



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application Number : 10/673,886 Confirmation No. 7856
Applicants : Jean-Louis ESCARY
Filed : September 30, 2003
Title : NEW POLYNUCLEOTIDES AND POLYPEPTIDES OF THE
IFN α -21 GENE
TC/Art Unit : 1634
Examiner: : Bruce D. HISSONG
Docket No. : 60711.000023
Customer No. : **21967**

SUBMISSION OF CERTIFIED COPY OF PRIORITY DOCUMENT

MAIL STOP AMENDMENT
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Sir:

Applicant respectfully submits the certified copy of French Patent Application No. 01/04404, filed March 30, 2001 in connection with the above-identified patent application.

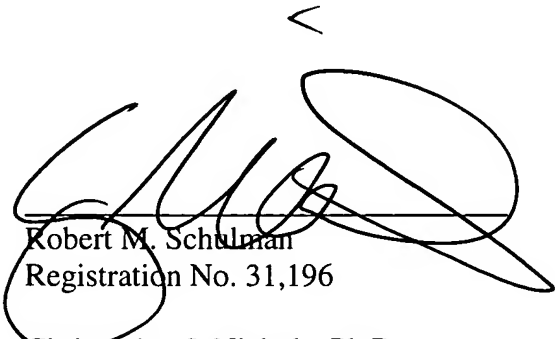
CONCLUSION

No fee is believed due as a result of this submission. However, if a fee is due upon the filing of this priority document, please charge such fee to the undersigned's Deposit Account No. 50-0206.

Respectfully submitted,
HUNTON & WILLIAMS LLP

Dated: June 12, 2008

By:



Robert M. Schulman
Registration No. 31,196

Christopher J. Nichols, Ph.D.
Registration No. 55,984

HUNTON & WILLIAMS LLP
Intellectual Property Department
1900 K Street, N.W., Suite 1200
Washington, DC 20006-1109
(202) 955-1500 (telephone)
(202) 778-2201 (facsimile)



(IFN x - 21)

Brevet d'invention

Certificat d'utilité

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 08 AVR. 2008

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

A handwritten signature in black ink, appearing to read 'M+Planche', enclosed within a large, horizontal oval shape.

Martine PLANCHE

| | | | |
|---|--|--|--|
| <div style="border: 1px solid black; display: inline-block; padding: 2px;">Réservé à l'INPI</div> | | <div style="border: 1px solid black; padding: 5px;"> 1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE RINUY, SANTARELLI 14, avenue de la Grande Armée 75017 PARIS </div> | |
| <div style="border: 1px solid black; padding: 5px;"> REMISE DES PIÈCES DATE 30 MARS 2001 LIEU 75 INPI PARIS N° D'ENREGISTREMENT NATIONAL ATTRIBUÉ PAR L'INPI 0104404 DATE DE DÉPÔT ATTRIBUÉE PAR L'INPI 30 MARS 2001 </div> | | <div style="border: 1px solid black; padding: 5px;"> Vos références pour ce dossier <i>(facultatif)</i> BIF022965/FR </div> | |
| <div style="border: 1px solid black; padding: 5px;"> Confirmation d'un dépôt par télécopie <input type="checkbox"/> N° attribué par l'INPI à la télécopie </div> | | <div style="border: 1px solid black; padding: 5px;"> 2 NATURE DE LA DEMANDE Demande de brevet <input checked="" type="checkbox"/> Demande de certificat d'utilité <input type="checkbox"/> Demande divisionnaire <input type="checkbox"/> <i>Demande de brevet initiale</i> N° _____ Date ____/____/____ <i>ou demande de certificat d'utilité initiale</i> N° _____ Date ____/____/____ Transformation d'une demande de brevet européen <i>Demande de brevet initiale</i> <input type="checkbox"/> N° _____ Date ____/____/____ </div> | |
| <div style="border: 1px solid black; padding: 10px;"> 3 TITRE DE L'INVENTION (200 caractères ou espaces maximum) Nouveaux polynucléotides comportant des polymorphismes de type SNP fonctionnels dans la séquence nucléotidique du gène IFNα-21 ainsi que de nouveaux polypeptides codés par ces polynucléotides et leurs utilisations thérapeutiques. </div> | | | |
| <div style="border: 1px solid black; padding: 5px;"> 4 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE FRANÇAISE </div> | | <div style="border: 1px solid black; padding: 5px;"> Pays ou organisation _____ N° _____ Date ____/____/____ Pays ou organisation _____ N° _____ Date ____/____/____ Pays ou organisation _____ N° _____ Date ____/____/____ <input type="checkbox"/> S'il y a d'autres priorités, cochez la case et utilisez l'imprimé «Suite» </div> | |
| <div style="border: 1px solid black; padding: 5px;"> 5 DEMANDEUR </div> | | <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> S'il y a d'autres demandeurs, cochez la case et utilisez l'imprimé «Suite» </div> | |
| <div style="border: 1px solid black; padding: 5px;"> Nom ou dénomination sociale Prénoms Forme juridique N° SIREN Code APE-NAF Adresse Rue Code postal et ville Pays Nationalité N° de téléphone <i>(facultatif)</i> N° de télécopie <i>(facultatif)</i> Adresse électronique <i>(facultatif)</i> </div> | | <div style="border: 1px solid black; padding: 5px;"> GenOdyssee Société Anonyme Parc d'Affaires Technopolis, 3 avenue du Canada, Bat Alpha, B.P. 810, LES ULIS 91974 COURTABOEUF FRANCE FRANÇAISE </div> | |



BREVET D'INVENTION CERTIFICAT D'UTILITÉ

REQUÊTE EN DÉLIVRANCE 2/2

| | | | |
|--|--|--|-------------------|
| REMISE DES PIÈCES DATE 30 MARS 2001 LIEU 75 INPI PARIS N° D'ENREGISTREMENT 0104404 NATIONAL ATTRIBUÉ PAR L'INPI | | Réservé à l'INPI | DB 540 W / 190600 |
| Vos références pour ce dossier : <i>(facultatif)</i> | | BIF022965/FR | |
| 6 MANDATAIRE Nom Prénom Cabinet ou Société N° de pouvoir permanent et/ou de lien contractuel Adresse Rue Code postal et ville N° de téléphone <i>(facultatif)</i> N° de télécopie <i>(facultatif)</i> Adresse électronique <i>(facultatif)</i> | | RINUY, SANTARELLI 14 AVENUE DE LA GRANDE ARMÉE 75017 PARIS 01 40 55 43 43 | |
| 7 INVENTEUR (S) | | | |
| Les inventeurs sont les demandeurs | | <input type="checkbox"/> Oui <input checked="" type="checkbox"/> Non Dans ce cas fournir une désignation d'inventeur(s) séparée | |
| 8 RAPPORT DE RECHERCHE | | Uniquement pour une demande de brevet (y compris division et transformation) | |
| Établissement immédiat ou établissement différé | | <input checked="" type="checkbox"/> <input type="checkbox"/> | |
| Paiement échelonné de la redevance | | Paiement en deux versements, uniquement pour les personnes physiques <input type="checkbox"/> Oui <input type="checkbox"/> Non | |
| 9 RÉDUCTION DU TAUX DES REDEVANCES | | Uniquement pour les personnes physiques <input type="checkbox"/> Requête pour la première fois pour cette invention <i>(joindre un avis de non-imposition)</i> <input type="checkbox"/> Requête antérieurement à ce dépôt <i>(joindre une copie de la décision d'admission pour cette invention ou indiquer sa référence) :</i> | |
| Si vous avez utilisé l'imprimé «Suite», indiquez le nombre de pages jointes | | | |
| 10 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE (Nom et qualité du signataire) | | VISA DE LA PRÉFECTURE OU DE L'INPI | |
| Thierry CAEN N°98.0600 RINUY, SANTARELLI | | | |

5 La présente invention concerne de nouveaux polynucléotides comportant des polymorphismes de type SNP fonctionnels dans la séquence nucléotidique du gène IFN α -21 ainsi que de nouveaux polypeptides codés par ces polynucléotides et leurs utilisations thérapeutiques.

10 ART ANTERIEUR

 Le gène de l'interféron alpha-21, ci après IFN α -21 est décrit dans les publications :

- Goeddel, D. V., Leung, D. W; "The structure of eight distinct cloned human leukocyte interferon cDNAs"; Nature 290 (5801), 20-26 (1981), et
- 15 - Olopade OI., Bohlander SK.; "Mapping of the shortest region of overlap of deletions of the short arm of chromosome 9 associated with human neoplasia"; Genomics 14 (2), 437-443 (1992).

 La séquence nucléotidique de ce gène est accessible dans la section HTG de la base de donnée GenBank sous le numéro d'accès

20 AC009445.

 La séquence de l'ARN messager de l'IFN α -21 est mentionné dans la base de données du NCBI, sous le code d'accès NM_002175.

 L'IFN α -21 est un gène ayant une homologie structurelle et fonctionnelle proche des interférons alpha humains (l'IFN α), en particulier de

25 l'IFN α -2.

 Les IFN α sont connus pour leurs effets anti-prolifératifs cellulaires et leurs implications dans les réponses anti-virales et anti-parasitaires.

 Les IFN α sont connus pour inhiber l'expression de plusieurs autres cytokines au niveau des cellules hématopoïétiques souches, ainsi que la

30 prolifération cellulaire de certaines tumeurs cancéreuses ou non.

 Les IFN α sont également connus pour réduire l'expression des récepteurs à l'EGF dans les carcinomes rénaux, pour inhiber l'expression de



certaines gènes mitochondriaux, pour inhiber la prolifération des fibroblastes, des monocytes et des lymphocytes B, notamment in vitro et pour bloquer la synthèse des anticorps dans les lymphocytes B.

Les IFN α sont aussi connus pour induire l'expression d'antigènes
5 spécifiques de tumeurs à la surface de cellules tumorales et également pour induire les gènes placés sous le contrôle de régions promotrices de type ISRE (Interferon-Stimulated Response Element) en agissant sur les facteurs de transcription spécifiques de ces ISRE.

Il est connu que les IFN α sont impliqués dans différents
10 dérèglements et/ou maladies humaines, tels que les différents cancers comme, par exemple, les carcinomes, les mélanomes, les lymphomes, les myélomes, les leucémies et les cancers du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que celles qui ne sont pas liées au système immunitaire comme, par exemple, l'obésité, les
15 maladies infectieuses en particulier les infections virales comme les hépatites B et C et le SIDA, les pneumonies, les colites ulcéraives, les maladies du système nerveux central comme, par exemple, la schizophrénie et la dépression, le rejet de greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie en particulier chez les patients dialysés, les allergies,
20 l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les maladies et les désordres auto-immuns, les désordres gastro-intestinaux ou encore les désordres liés aux traitements par chimiothérapie.

Les IFN α sont particulièrement utilisés pour le traitement des
25 hépatites chroniques B et C, des leucémies telles que la leucémie à trichloroleucocytes et la leucémie myéloïde chronique, des myélomes multiples, des lymphomes folliculaires, de tumeurs carcinoïdes, de mélanomes malins, de carcinomes rénaux métastasés ainsi que des tumeurs qui apparaissent à la suite d'un déficit immunitaire, tel que le sarcome de Kaposi dans le cas du
30 SIDA.

Il est également connus que les IFN α sont aussi efficaces contre les tumeurs cancéreuses ou non. Les IFN α sont reconnus par la FDA (Food

and Drug Administration) pour le traitement des verrues génitales ou vénériennes.

Plus particulièrement, l'IFN α -21 a pût être localisé par hybridation *in situ* dans des cerveaux de patients atteints de la maladie de Parkinson ou de la maladie d'Alzheimer.

Les cellules microgliales expriment en grande quantité l'IFN α -21, par rapport aux autres cellules.

Chez les patients atteints de maladie d'Alzheimer, la présence de l'IFN α -21 a pût être démontrée dans des neurones des lobes pariétaux, suggérant ainsi que l'IFN α -21 pourrait être impliqué dans le développement de cette pathologie (Voir Kawaguchi N, Yamada T, Yoshiyama Y. No To Shinkei. 1997 Jan.; 49(1) : 69-73).

Toutefois, l'IFN α -21, comme les IFN α de manière générale, présentent de nombreux effets secondaires lorsqu'ils sont employés dans des compositions pharmaceutiques, tels que des réactions d'hypersensibilité aiguë (urticaire, broncho-constriction, choc anaphylactique, etc.), des arythmies cardiaques, des hypotensions artérielles, des crises d'épilepsie, des troubles des fonctions thyroïdiennes ou des syndromes pseudo-grippaux (fièvres, sueurs, myalgies).

La demanderesse a trouvé de nouveaux polypeptides et nouveaux polynucléotides analogues à l'IFN α -21, ayant une fonctionnalité différente de l'IFN α -21 naturel, pouvant être utilisés notamment pour traiter ou prévenir les dérèglements ou maladies mentionnés précédemment et éviter tout ou partie des inconvénients qui leurs sont liés.

25

L'INVENTION

L'invention a pour principal objet de nouveaux polynucléotides qui diffèrent de la séquence nucléotidique du gène sauvage de référence de l'IFN α -21, en ce qu'ils comportent un ou plusieurs polymorphismes de type SNP (Single Nucleotide Polymorphism) fonctionnels.

30

La demanderesse a en effet découvert 5 polymorphismes de type SNP fonctionnels dans la séquence nucléotidique du gène sauvage de

référence de l'IFN α -21, correspondant à la séquence nucléotidique ID SEQ N°1.

La séquence nucléotidique ID SEQ N°1 comporte 2001 nucléotides.

5 Ces 5 polymorphismes de type SNP fonctionnels sont présentés ci-dessous :

1) Le nucléotide cytosine (c) en position 973 de la séquence nucléotidique ID SEQ N° 1 du gène sauvage de référence IFN α -21 est muté en adénine (a).

10 Ce polymorphisme de type SNP est appelé ci-après c973a.

2) Le nucléotide guanine (g) en position 1011 de la séquence nucléotidique ID SEQ N° 1 du gène sauvage de référence IFN α -21 est muté en cytosine (c).

Ce polymorphisme de type SNP est appelé ci-après g1011c.

15 3) Le nucléotide thymine (t) en position 1049 de la séquence nucléotidique ID SEQ N° 1 du gène sauvage de référence IFN α -21 est muté en adénine (a).

Ce polymorphisme de type SNP est appelé ci-après t1049a.

20 4) Le nucléotide thymine (t) en position 1155 de la séquence nucléotidique ID SEQ N° 1 du gène sauvage de référence IFN α -21 est muté en adénine (a).

Ce polymorphisme de type SNP est appelé ci-après t1155a.

25 5) Le nucléotide adénine (a) en position 1204 de la séquence nucléotidique ID SEQ N° 1 du gène sauvage de référence IFN α -21 est muté en guanine (g).

Ce polymorphisme de type SNP est appelé ci-après a1204g.

30 Ces polymorphismes de type SNP fonctionnels ont été identifiés par la demanderesse au moyen du procédé de détermination décrit dans sa demande de brevet FR 00 22894, intitulée "Procédé de détermination d'un ou plusieurs polymorphisme(s) fonctionnel(s) dans la séquence nucléotidique d'un gène "candidat" fonctionnel présélectionné et ses applications" et déposée le 6 décembre 2000, citée ici à titre de référence.

Le procédé décrit dans cette demande de brevet permet l'identification d'un (ou plusieurs) polymorphisme(s) de type SNP fonctionnel(s) préexistant(s) dans au moins un individu constitutif d'une population aléatoire d'individus.

5 Dans le cadre de la présente invention, un fragment de la séquence nucléotidique du gène IFN α -21, comprenant la séquence codante, a été isolé chez 278 individus, dans la population d'individus choisis de manière aléatoire.

Un séquençage de ces fragments a ensuite été réalisé sur
10 certains de ces échantillons présentant un profil hétéroduplex après analyse par DHPLC (voir brevet FR 00 22894). On a alors comparé le fragment ainsi séquencé à la séquence nucléotidique du fragment du gène IFN α -21 sauvage de référence et identifier tout polymorphisme de type SNP fonctionnel.

Ainsi les polymorphismes de type SNP fonctionnels selon
15 l'invention sont naturels et sont présents chez certains individus de la population mondiale.

Un génotypage des polynucléotides conformes à l'invention a été effectué de façon à déterminer la fréquence allélique de ces polynucléotides dans une population.

20 Des exemples de tels génotypages sont donnés, ci-après, dans la partie expérimentale.

La séquence nucléotidique ID SEQ N°1 du gène IFN α -21 humain sauvage de référence comporte une séquence codante de 570 nucléotides, du nucléotide 670 (codon start) au nucléotide 1239 (codon stop).

25 Le gène IFN α -21 humain sauvage de référence code pour une protéine immature de 189 acides aminés, correspondant à la séquence d'acides aminés ID SEQ N°2, qui sera convertit en protéine mature de 166 acides aminés (de l'acide aminé 24 à 189), par clivage du peptide signal qui comprend les 23 premiers acides aminés.

30 Chacun des polymorphismes de type SNP de l'invention : c973a, g1011c, t1049a, t1155a et a1204g, entraînent des modifications de la protéine codée par la séquence nucléotidique du gène IFN α -21.

Ces modifications dans la séquence d'acides aminés sont les suivantes:

1) Le polymorphisme de type SNP c973a entraîne une mutation de l'acide aminé glutamine (Q) en position 102 dans la protéine immature du gène IFN α -21, correspondant à la séquence d'acides aminés ID SEQ N°2, en lysine (K) et en position 79 de la protéine mature.

Dans la description de la présente invention, on appellera indifféremment Q79K et Q102K la mutation codée par ce polymorphisme de type SNP de l'invention selon que l'on se réfère respectivement à la protéine mature ou à la protéine immature.

2) Le polymorphisme de type SNP g1011c entraîne une mutation de l'acide aminé glutamine (Q) en position 114 dans la protéine immature du gène IFN α -21, correspondant à la séquence d'acides aminés ID SEQ N°2, en histidine (H) et en position 91 de la protéine mature.

Dans la description de la présente invention, on appellera indifféremment Q91H et Q114H la mutation codée par ce polymorphisme de type SNP de l'invention selon que l'on se réfère respectivement à la protéine mature ou à la protéine immature.

3) Le polymorphisme de type SNP t1049a entraîne une mutation de l'acide aminé valine (V) en position 127 dans la protéine immature du gène IFN α -21, correspondant à la séquence d'acides aminés ID SEQ N°2, en acide aspartique (D) et en position 104 de la protéine mature.

Dans la description de la présente invention, on appellera indifféremment V104D et V127D la mutation codée par ce polymorphisme de type SNP de l'invention selon que l'on se réfère respectivement à la protéine mature ou à la protéine immature.

4) Le polymorphisme de type SNP t1155a entraîne une mutation de l'acide aminé cystéine en position 162 dans la protéine immature du gène IFN α -21, correspondant à la séquence d'acides aminés ID SEQ N°2, en codon stop et en position 139 de la protéine mature.

Dans la description de la présente invention, on appellera indifféremment C139stop et C162stop la mutation codée par ce polymorphisme

de type SNP de l'invention selon que l'on se réfère respectivement à la protéine mature ou à la protéine immature.

5 5) Le polymorphisme de type SNP a1204g entraîne une mutation de l'acide aminé lysine (K) en position 179 dans la protéine immature du gène IFN α -21, correspondant à la séquence d'acides aminés ID. SEQ N°2, en glutamate (E) et en position 156 de la protéine mature.

Dans la description de la présente invention, on appellera indifféremment K156E et K179E la mutation codée par ce polymorphisme de type SNP de l'invention selon que l'on se réfère respectivement à la protéine
10 mature ou à la protéine immature.

Les polypeptides conformes à l'invention comportant l'un au moins des polymorphismes de type SNP décrit ci-dessus entraînent des modifications de la conformation spatiale ceux par rapport au polypeptide de l'IFN α -21 naturel.

15 Ces modifications peuvent être observées par modélisation moléculaire bio-informatique, selon des méthodes bien connues de l'homme du métier mettant en œuvre, par exemple, les outils de modélisation *de novo* (par exemple, SEQFOLD/MSI), d'homologie (par exemple, MODELER/MSI), de minimisation des champs de force (par exemple, DISCOVER, DELPHI/MSI)
20 et/ou de dynamique moléculaire (par exemple, CFF/MSI).

Des exemples de telles modélisations sont donnés ci-après dans la partie expérimentale.

1) La modélisation bio-informatique permet d'observer que la mutation Q102K sur la protéine mutée mature entraîne un déplacement de
25 l'hélice dite A de l'IFN α -21 naturel due à des perturbations de liaisons hydrogène, comme il est observé sur les Figures 1 et 2.

En effet, les liaisons hydrogènes de l'oxygène de la chaîne latérale de Q102 avec le résidu acide de E106 et avec l'hélice dite C de l'IFN α -21 naturel disparaissent dans la protéine mutée K102.

30 Ainsi, la protéine mutée possède une conformation tridimensionnelle très différente de la protéine codée par le gène sauvage de IFN α -21 de référence.

La modélisation bio-informatique permet ainsi de prévoir que la présence de la lysine en position 102 entraîne une modification significative de la structure et de la fonction de l'IFN α -21.

2) La modélisation bio-informatique permet d'observer que la
5 mutation Q114H sur la protéine mutée mature entraîne un déplacement de l'hélice dite C au niveau du point de mutation, comme il est observé sur les Figures 3 et 4.

Plusieurs liaisons hydrogènes et ponts salins apparaissent, entre
autres, entre les chaînes latérales des acides aminés H114 et D99,
10 occasionnant une rigidification de l'hélice.

Ainsi, la protéine mutée possède une conformation tridimensionnelle différente de la protéine codée par le gène sauvage de IFN α -21 de référence.

La modélisation bio-informatique permet ainsi de prévoir que la
15 présence de l'histidine en position 114 entraîne une modification significative de la structure et de la fonction de l'IFN α -21.

3) La modélisation bio-informatique permet d'observer que la
mutation V127D sur la protéine mutée mature entraîne des modifications
structurales de l'hélice A au niveau du point de mutation, comme il est observé
20 sur les Figures 5 et 6.

Il a été observé que ces modifications spatiales affectent des résidus impliqués dans la liaison de IFN α -21 à son récepteur.

La modélisation bio-informatique permet ainsi de prévoir que la
présence de l'acide aspartique en position 127 entraîne une modification
25 significative de la conformation tridimensionnelle et de la fonction de l'IFN α -21.

4) La modélisation bio-informatique permet d'observer que la
mutation C162stop entraîne un arrêt prématuré de la traduction de la protéine
avec pour conséquence la disparition d'un fragment polypeptidique qui participe
à la dernière hélice, dite hélice E, de l'IFN α -21 naturel, comme il est observé
30 sur la Figure 7.

L'hélice E est primordiale pour l'accrochage de l'IFN α -21 à son récepteur.

L'absence de cette hélice E dans la protéine mutée entraîne un mauvais repliement de la protéine mutée et induit ainsi une modification de la conformation tridimensionnelle dans laquelle le cœur hydrophobe de cette protéine se trouve en contact avec le milieu extérieur hydrophile.

5 La protéine doit ainsi modifier sa conformation tridimensionnelle de façon à couvrir son cœur hydrophobe par des résidus hydrophiles, pour éviter le contact avec milieu extérieur hydrophile.

En conséquence, la protéine mutée possède une conformation tridimensionnelle très différente de la protéine codée par le gène sauvage de
10 IFN α -21 de référence.

La modélisation bio-informatique permet ainsi de prévoir que l'arrêt prématuré de la protéine en position 162 entraîne une modification significative de la structure et de la fonction de l'IFN α -21.

5) La modélisation bio-informatique permet d'observer que la
15 mutation K179E sur la protéine mutée mature entraîne un dépliement de l'hélice E et une modification la forme de la boucle C-terminale, comme il est observé sur les Figures 8 et 9.

Cette mutation s'accompagne d'une augmentation du réseau de liaisons hydrogènes avec la création d'un pont salin entre E179 (E156 dans la
20 protéine mature) et R184 (R161 dans la protéine mature).

Ces modifications entraînent ainsi une rigidification de la structure de l'IFN α -21 dans cette région.

La région protéique affectée par cette mutation est connue pour être impliquée dans son activité anti-virale.

25 Par conséquent, cette mutation permet de prévoir que l'activité anti-virale de l'IFN α -21 K179E sera fortement perturbée et que l'acide glutamique en position 179 entraîne une modification de la structure et de la fonction de l'IFN α -21.

La détermination de la fonctionnalité des polypeptides de
30 l'invention peut également être effectuée par un test de l'activité biologique en présence de polypeptides conformes à l'invention et d'un polypeptide codé par le gène sauvage de référence de l'IFN α -21, par exemple, par le test d'activité



anti-proliférative de l'IFN α -21 sur la lignée humaine tumorale Daudi du lymphome Burkitt (Journ. Biol. Chem.; Vol. 275, Issue 51, 40425-40433, December 22, 2000; "New structural and Functional Aspects of the type I Interferon-Receptor interaction revealed by comprehensible mutational analysis of the binding interface"; Pielher et al..

L'invention a aussi pour objet la mise en évidence et l'utilisation de molécules thérapeutiques obtenues à partir de l'information dérivant de polynucléotides et de polypeptides conformes à l'invention, notamment pour la prévention et le traitement des différents dérèglements et/ou maladies humaines.

La Figure 1 représente d'une part la modélisation de la protéine conforme à l'invention comportant le polymorphisme de type SNP c973a et d'autre part la protéine IFN α -21 naturel.

La Figure 2 représente la modélisation de la partie inférieure de chacune des protéines représentées sur la Figure 1.

Sur ces deux Figures, les zones blanches représentent la structure de la protéine IFN α -21 naturel et les zones noires représentent la structure de la protéine conforme à l'invention comportant le polymorphisme de type SNP c973a.

La Figure 10 représente le résultat du génotypage du polymorphisme de type SNP c973a dans une population d'individus.

Sur cette figure, les abscisses représentent la valeur mp du filtre Tamra (ddTTP) et les ordonnées représentent la valeur mp du filtre R-110 (ddGTP).

En haut à droite, le groupe hétérozygote GT contient 4 individus.
En haut à gauche, le groupe homozygote GG contient 233 individus.

En bas à gauche, le groupe d'individus contient 7 individus blancs et 2 individus non génotypés.

La Figure 3 représente la modélisation de la protéine conforme à l'invention comportant le polymorphisme de type SNP g1011c et de la protéine IFN α -21 naturel.

La Figure 4 représente la modélisation de la partie gauche de chacune des protéines représentées sur la Figure 3.

Sur ces deux Figures, les zones sombres représentent la structure de la protéine IFN α -21 naturel et les zones claires représentent la structure de la protéine conforme à l'invention comportant le polymorphisme de type SNP g1011c.

La Figure 11 représente le résultat du génotypage du polymorphisme de type SNP g1011c dans une population d'individus.

Sur cette figure, les abscisses représentent la valeur mp du filtre Tamra (ddCTP) et les ordonnées représentent la valeur mp du filtre R-110 (ddGTP).

En haut au milieu, le groupe hétérozygote CG contient 1 individu.

En bas à droite, le groupe homozygote CC contient 236 individus.

En bas à gauche, le groupe d'individus contient 7 individus blancs et 1 individu non génotypé.

La Figure 5 représente la modélisation de la protéine conforme à l'invention comportant le polymorphisme de type SNP t1049a et de la protéine IFN α -21 naturel.

La Figure 6 représente la modélisation de la partie supérieure de chacune des protéines représentées sur la Figure 5.

Sur ces deux Figures, les zones noires représentent la structure de la protéine IFN α -21 naturel et les zones blanches représentent la structure de la protéine conforme à l'invention comportant le polymorphisme de type SNP t1049a.

La Figure 12 représente le résultat du génotypage du polymorphisme de type SNP t1049a dans une population d'individus.

Sur cette figure, les abscisses représentent la valeur mp du filtre Tamra (ddTTP) et les ordonnées représentent la valeur mp du filtre R-110 (ddATP).

En haut à droite, le groupe hétérozygote AT contient 1 individu.

En haut à gauche, le groupe homozygote AA contient 236 individus.

En bas à gauche, le groupe d'individus contient 7 individus blancs et 2 individus non génotypé.

La Figure 7 représente la modélisation de la protéine conforme à l'invention comportant le polymorphisme de type SNP t1155a (à droite de la Figure) et de la protéine IFN α -21 naturel (à gauche de la Figure). Au centre de la Figure, les deux modélisations moléculaires sont superposées.

Sur cette Figure, les zones claires représentent la structure de la protéine IFN α -21 naturel et les zones sombres représentent la structure de la protéine conforme à l'invention comportant le polymorphisme de type SNP t1155a.

La Figure 13 représente le résultat du génotypage du polymorphisme de type SNP t1155a dans une population d'individus.

Sur cette figure, les abscisses représentent la valeur mp du filtre Tamra (ddTTP) et les ordonnées représentent la valeur mp du filtre R-110 (ddATP).

En haut à droite, le groupe hétérozygote AT contient 1 individu.

En haut à gauche, le groupe homozygote AA contient 235 individus.

En bas à gauche, le groupe d'individus contient 7 individus blancs et 2 individus non génotypés.

La Figure 8 représente la modélisation de la protéine conforme à l'invention comportant le polymorphisme de type SNP a1204g et de la protéine IFN α -21 naturel.

La Figure 9 représente la modélisation de la partie supérieure gauche de chacune des protéines représentées sur la Figure 1.

Sur ces deux Figures, les zones sombres représentent la structure de la protéine IFN α -21 naturel et les zones claires représentent la structure de la protéine conforme à l'invention comportant le polymorphisme de type SNP a1204g.

La Figure 14 représente le résultat du génotypage du polymorphisme de type SNP a1204g dans une population d'individus.

Sur cette figure, les abscisses représentent la valeur m_p du filtre Tamra (ddATP) et les ordonnées représentent la valeur m_p du filtre R-110 (ddGTP).

- 5 En haut à droite, le groupe hétérozygote GA contient 13 individus.
 En bas à droite, le groupe homozygote AA contient 225 individus.
 En bas à gauche, le groupe d'individus contient 7 individus blancs et 1 individu non génotypé.

DESCRIPTION DETAILLEE DE L'INVENTION

10 Définitions

On entend par "séquence nucléotidique du gène sauvage de référence", la séquence nucléotidique ID SEQ N° 1 du gène IFN α -21 humain, représenté ci-après.

- 15 Cette séquence nucléotidique est accessible dans une large
 séquence répertoriée de la section HTG de la base de donnée GenBank sous le numéro d'accès AC009445.

On entend par "IFN α -21 naturel" le polypeptide codé par la séquence nucléotidique du gène sauvage de référence. La protéine immature de l'IFN α -21 naturel correspond à la séquence d'acides aminés ID SEQ N°2.

- 20 On entend par "polynucléotide", un polyribonucléotide ou un
 polydésoxyribonucléotide qui peut être un ADN ou un ARN modifié ou non.

- Le terme polynucléotide inclut, par exemple, un ADN simple brin ou double brin, un ADN composé d'un mélange d'une ou plusieurs région(s) simple brin et d'une ou plusieurs région(s) double brins, un ARN simple brin ou
 25 double brin et un ARN composé d'un mélange d'une ou plusieurs région(s) simple brin et d'une ou plusieurs région(s) double brins. Le terme polynucléotide peut aussi comprendre un ARN et/ou un ADN comprenant une ou plusieurs régions triple brins. On entend également par polynucléotide les ADNs et ARNs contenant une ou plusieurs bases modifiées de façon à avoir un squelette
 30 modifié pour la stabilité ou pour d'autres raisons. On entend par base modifiée, par exemple, les bases inhabituelles telles que l'inosine.



On entend par "polypeptide", un peptide, un oligopeptide, un oligomère ou une protéine comprenant au moins deux acides aminés joints l'un à l'autre par une liaison peptidique normale ou modifiée, comme dans le cas des peptides isostères, par exemple.

5 Un polypeptide peut être composé d'autres acides aminés que les 20 acides aminés codés par les gènes humains. Un polypeptide peut également être composé d'acides aminés modifiés par des processus naturels, tel que le processus de maturation post-traductionnel ou par des procédés chimiques, qui sont bien connus de l'homme du métier. De telles modifications
10 sont bien détaillées dans la littérature. Ces modifications peuvent apparaître n'importe où dans le polypeptide : dans le squelette peptidique, dans la chaîne d'acides aminés ou encore aux extrémités carboxy- ou amino-terminales.

Un polypeptide peut être ramifié suite à une ubiquitination ou être cyclique avec ou sans ramification. Ce type de modifications peut être le
15 résultat de processus de post-translation naturel ou synthétique, qui sont bien connus de l'homme du métier.

On entend, par exemple, par modifications d'un polypeptide, l'acétylation, l'acylation, l'ADP-ribosylation, l'amidation, la fixation covalente de flavine, la fixation covalente d'un hème, la fixation covalente d'un nucléotide ou
20 d'un dérivé nucléotidique, la fixation covalente d'un lipide ou d'un dérivé lipidique, la fixation covalente d'un phosphatidylinositol, la réticulation covalente ou non-covalente, la cyclisation, la formation de pont disulfure, la déméthylation, la formation de cystéine, la formation de pyroglutamate, la formylation, la gamma-carboxylation, la glycosylation, la formation d'ancre au
25 GPI, l'hydroxylation, l'iodisation, la méthylation, la myristoylation, l'oxydation, le processus protéolytique, la phosphorylation, la prénylation, la racémisation, la sénéloylation, la sulfatation, l'addition d'acides aminés telle que l'arginylation ou l'ubiquitination. De telles modifications sont bien détaillées dans la littérature :
PROTEINS-STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E.
30 Creighton, New York, 1993, POST-TRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983, Seifter et al. "Analysis for protein modifications and nonprotein

cofactors", Meth. Enzymol. (1990) 182 :626-646 et Rattan et al. "Protein Synthesis : Post-translational Modifications and Aging", Ann NY Acad. Sci. (1992) 663 : 48-62.

On entend par "polynucléotide isolé" ou "polypeptide isolé" un polynucléotide ou un polypeptide tel que défini précédemment qui est isolé du corps humain ou autrement produit par un procédé technique.

On entend par "identité", la mesure d'identité d'une séquence nucléotidique ou polypeptidique. D'une manière générale, les séquences que l'on veut comparer sont alignées de façon à obtenir la plus grande identité entre ces séquences.

L'identité est un terme bien connu de l'homme du métier et de la littérature. Voir COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., Ed., Oxford University Press, New York, 1998; BIOCOMPUTING INFORMATICS AND GENOME PROJECT, Smith, D.W., Ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M. and Griffin H.G., Ed, Humana Press, New Jersey, 1994; et SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987.

Les méthodes communément employées pour déterminer l'identité et la similarité entre deux séquences sont également bien décrites dans la littérature. Voir GUIDE TO HUGE COMPUTER, Martin J. Bishop, Ed, Academic Press, San Diego, 1994, et Carillo H. and Lipton D., Siam J Applied Math (1988) 48 : 1073.

Un polynucléotide ayant, par exemple, une identité d'au moins 95 % avec la séquence nucléotidique ID SEQ N° 1 est un polynucléotide qui comporte au plus 5 points de mutation sur 100 nucléotides, par rapport à ladite séquence.

Ces points de mutation peuvent être une (ou plusieurs) substitution(s), addition(s) et/ou délétion(s) d'un (ou plusieurs) nucléotide(s).

De même, un polypeptide ayant, par exemple, une identité d'au moins 95 % avec la séquence d'acides aminés ID SEQ N° 2 est un polypeptide qui comporte au plus 5 points de mutation sur 100 acides aminés, par rapport à ladite séquence.



Ces points de mutation peuvent être une (ou plusieurs) substitution(s), addition(s) et/ou délétion(s) d'un (ou plusieurs) acide(s) aminé(s).

5 Les polynucléotides et les polypeptides conformes à l'invention qui ne sont pas totalement identiques avec respectivement la séquence nucléotidique ID SEQ N°1 ou la séquence d'acides aminés ID SEQ N°2, étant entendu que ces séquences comportent le polymorphisme de type SNP de l'invention, sont considérés comme des variants de ces séquences.

10 Habituellement un polynucléotide conforme à l'invention possède la même, ou pratiquement la même activité biologique, que la séquence nucléotidique ID SEQ N°1 comportant au moins l'un des polymorphismes de type SNP suivants : c973a, g1011c, t1049a, t1155a, a1204g.

15 De même, habituellement un polypeptide conforme à l'invention possède la même, ou pratiquement la même activité biologique que la séquence d'acides aminés ID SEQ N°2 comportant au moins l'un des polymorphismes de type SNP suivants : Q102K, Q114H, V127D, C162stop, K179E.

Un variant, selon l'invention, peut être obtenu, par exemple, par mutagenèse dirigée ou par synthèse directe.

20 On entend par "polymorphisme de type SNP", toute variation naturelle d'une base dans une séquence nucléotidique.

25 On entend par "polymorphisme de type SNP fonctionnel", un polymorphisme de type SNP qui a pour conséquence de modifier la fonctionnalité d'un polynucléotide ou d'un polypeptide codé par ce polynucléotide.

On entend par "fonctionnalité", l'activité biologique d'un polypeptide ou d'un polynucléotide codant pour ce polypeptide.

30 La fonctionnalité d'un polypeptide conforme à l'invention ou d'un polynucléotide codant pour ce polypeptide peut consister en une augmentation, une diminution ou une suppression de l'activité biologique, voire un changement de nature total ou partiel de l'activité biologique soit du polypeptide

codé par la séquence nucléotidique du gène sauvage de référence soit de cette dernière séquence nucléotidique.

L'activité biologique peut, notamment, être liée à l'affinité ou à l'absence d'affinité du polypeptide vis-à-vis d'un ligand, tel qu'un récepteur.

5

Polynucléotide

La présente invention a pour premier objet un polynucléotide isolé comprenant :

- a) une séquence nucléotidique ayant au moins 80 % d'identité, 10 préférentiellement au moins 90 % d'identité, plus préférentiellement au moins 95 % d'identité et encore plus préférentiellement au moins 99 % d'identité avec la séquence ID SEQ N° 1 ou sa séquence codante (du nucléotide 670 au nucléotide 1239),
- sous réserve que cette séquence nucléotidique comporte au moins l'un des 15 polymorphismes de type SNP suivants :
- c973a,
 - g1011c,
 - t1049a,
 - t1155a,
 - 20 - a1204g; ou
- b) une séquence nucléotidique complémentaire à une séquence nucléotidique sous a).

Le polynucléotide conforme à l'invention peut ainsi comporter un ou plusieurs polymorphisme(s) de type SNP fonctionnel(s) défini(s) 25 précédemment, comme par exemple l'association des polymorphismes g1011c et t1049a.

La présente invention concerne également un polynucléotide isolé comprenant :

- a) la séquence nucléotidique ID SEQ N° 1 ou sa séquence 30 codante sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP suivants :
- c973a,

- g1011c,
- t1049a,
- t1155a,
- a1204g; ou

5 b) une séquence nucléotidique complémentaire à une séquence nucléotidique sous a).

Préférentiellement, le polynucléotide de l'invention consiste en la séquence ID SEQ N° 1 ou sa séquence codante, sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP
10 suivants : c973a, g1011c, t1049a, t1155a, a1204g; et sa séquence nucléotidique complémentaire.

Selon l'invention, le polynucléotide défini précédemment comporte préférentiellement un seul polymorphisme de type SNP choisi dans le groupe constitué par :

- 15 - c973a,
- g1011c,
 - t1049a,
 - t1155a, et
 - a1204g.

20 La présente invention a aussi pour objet un polynucléotide codant pour un polypeptide comprenant :

- a) la séquence d'acides aminés ID SEQ N° 2, ou
- b) la séquence d'acides aminés comprenant les acides aminés compris entre 24 et 189 de la séquence d'acides aminés ID SEQ N° 2,

25 sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

- Q102K,
- Q114H,
- V127D,

30 - C162stop,

- K179E.

Selon un objet préféré de l'invention, le polypeptide défini

précédemment comporte un seul polymorphisme de type SNP tel que défini ci-dessus.

Préférentiellement le polynucléotide conforme à l'invention comprend une molécule d'ADN ou d'ARN.

5 Le polynucléotide de l'invention peut être obtenu par les méthodes standards de synthèse d'ADN ou d'ARN.

Ce polynucléotide peut également être obtenu par mutagenèse dirigée à partir de la séquence nucléotidique du gène IFN α -21 en modifiant l'un au moins des nucléotides c, g, t, t et a par respectivement les nucléotides a, c, a, a et g en position 973, 1011, 1049, 1155 et 1204 sur la séquence ID SEQ N° 1.

Les procédés de mutagenèse dirigée qui peuvent ainsi être mis en œuvre sont bien connus de l'homme du métier. On peut notamment évoquer la publication de TA Kunkel en 1985 dans "Proc. Natl. Acad. Sci. USA" 82:488.

15 Le polynucléotide isolé peut également comprendre, par exemple, des séquences nucléotidiques codant pour des séquences d'acides aminés pre-, pro- ou pre-pro-protéine ou des séquences d'acides aminés marqueurs, comme l'hexa-histidine peptide.

Le polynucléotide de l'invention peut également être associé à des séquences nucléotidiques codant pour d'autres protéines ou fragments de protéines en vue d'obtenir des protéines de fusion ou à des fins de purification.

Le polynucléotide conforme à l'invention peut également comprendre des séquences nucléotidiques comme les séquences 5' et/ou 3' non-codantes, telles que, par exemple, des séquences transcrites ou non-transcrites, des séquences non-traduites, des séquences signal d'épissage, des séquences polyadénylées, des séquences de liaison avec des ribosomes ou encore des séquences qui stabilisent l'ARNm.

On définit comme séquence nucléotidique complémentaire à la séquence nucléotidique un polynucléotide qui peut être hybridé avec cette séquence nucléotidique, dans des conditions stringentes.

30 On entend généralement, mais pas nécessairement, par "conditions stringentes d'hybridation" les conditions chimiques qui permettent



une hybridation lorsque les séquences nucléotidiques ont une identité d'au moins 80 %, de préférence supérieure ou égale à 90 %, encore plus préférentiellement supérieure ou égale à 95 % et tout particulièrement supérieure ou égale à 99 %.

5 Dans le cadre de l'invention, lorsque les conditions stringentes d'hybridation permettent une hybridation des séquences nucléotidiques ayant une identité égale à 100 %, on considère que la séquence nucléotidique est strictement complémentaire à la séquence nucléotidique sous a), telle que décrite plus haut.

10 Les conditions stringentes peuvent être obtenues selon les méthodes bien connues de l'homme du métier et, par exemple, par une incubation des polynucléotides, à 42° C, dans une solution comprenant 50 % de formamide, 5xSSC (150 Mm de NaCl, 15 mM de trisodium citrate), 50 Mm de sodium phosphate (pH = 7,6), 5x Solution Denhardt, 10 % de dextran sulfate et
15 20 µg d'ADN de sperme de saumon dénaturé, suivi d'un lavage des filtres à 0,1x SSC, à 65° C.

Il est bien entendu, au sens de la présente invention, que la séquence nucléotidique complémentaire d'une séquence nucléotidique comportant l'un au moins des polymorphismes de type SNP conforme à
20 l'invention doit comporter le même polymorphisme de type SNP en anti-sens.

Ainsi, par exemple, si la séquence nucléotidique comporte le polymorphisme de type SNP a1204g, sa séquence nucléotidique complémentaire comporte toujours le nucléotide c en position 1204.

25 Identification, hybridation et/ou amplification d'un polynucléotide comportant le polymorphisme de type SNP

La présente invention a aussi pour objet l'utilisation de tout ou partie d'un polynucléotide défini précédemment, pour identifier, hybrider et/ou amplifier tout ou partie d'un polynucléotide consistant en la séquence ID SEQ
30 N° 1 ou sa séquence codante (du nucléotide 670 au nucléotide 1239), sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP suivants :



- c973a,
- g1011c,
- t1049a,
- t1155a,
- 5 - a1204g.

Génotypage et détermination de la fréquence du polymorphisme de type SNP

La présente invention a également pour objet l'utilisation de tout ou partie d'un polynucléotide conforme à l'invention comme outil de
10 génotypage.

La présente invention a aussi pour objet un procédé de détermination de la fréquence du polymorphisme de type SNP d'un polynucléotide conforme à l'invention dans lequel on procède à un génotypage chez un individu ou dans une population d'individus.

15 Au sens de l'invention, on définit le génotypage comme un procédé de détermination du génotype d'un individu ou d'une population d'individus.

Le polymorphisme de type SNP fonctionnel conforme à l'invention est préférentiellement génotypé dans une population d'individus.

20 On entend par "population d'individus", un groupe d'individus déterminés de façon aléatoire ou non. Ces individus peuvent être des humains, des animaux ou des plantes.

Lorsque ladite population d'individus est déterminée de façon non aléatoire, les individus peuvent être choisis selon leur ethnie ou selon leur
25 phénotype, notamment ceux qui sont atteints par les dérèglements et/ou maladies suivantes : les différents cancers comme, par exemple, les carcinomes, les mélanomes, les lymphomes, les myélomes, les tumeurs cancéreuses ou non, les leucémies et les cancers du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que
30 celles qui ne sont pas liées au système immunitaire comme, par exemple, l'obésité, les maladies infectieuses en particulier les infections virales comme les hépatites B et C et le SIDA, les pneumonies, les colites ulcéraives, les maladies du système nerveux central comme, par exemple, la maladie



d'Alzheimer, la schizophrénie et la dépression, le rejet de greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie en particulier chez les patients dialysés, les allergies, l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les maladies et les
5 désordres auto-immuns, les verrues génitales ou vénériennes, les désordres gastro-intestinaux ou encore les désordres liés aux traitements par chimiothérapie.

Le génotypage peut être effectué par un miniséquençage avec des ddNTPs chauds (2 ddNTPs différents marqués par des fluorophores
10 différents) et froids (2 ddNTPs non marqués), en liaison avec un lecteur de fluorescence polarisé. Le protocole de miniséquençage avec lecture de fluorescence polarisée (Technologie FP-TDI ou Fluorescence Polarization Template-direct Dye-Terminator Incorporation) est bien connu de l'homme du métier.

15 Il peut être réalisé sur un produit obtenu après amplification par réaction de polymérisation en chaîne (PCR) de l'ADN de chaque individu. Ce produit PCR est choisi pour couvrir la région génique du polynucléotide contenant le polymorphisme de type SNP étudié. Après la dernière étape dans le thermocycleur de la PCR, la plaque est alors placée sur un lecteur de
20 fluorescence polarisée pour la lecture des bases marquées en utilisant les filtres d'excitation et d'émission spécifique des fluorophores. Les valeurs d'intensité des bases marquées sont reportées sur un graphe.

Les amorces respectivement sens et antisens pour l'amplification PCR, dans le cas du polymorphisme de type SNP conforme à l'invention,
25 peuvent comprendre par exemple les séquences nucléotidiques ID SEQ N° 3 et ID SEQ N° 4.

Ces séquences nucléotidiques permettent d'amplifier un fragment d'une longueur de 696 nucléotides, du nucléotide 620 au nucléotide 1315, du gène de l'IFN α -21 (ces nucléotides étant numérotés en référence à la séquence
30 nucléotidique ID SEQ N° 1).

Une analyse statistique de la fréquence de chaque allèle (fréquence allélique) codé par le gène comportant le SNP dans la population

d'individus est alors effectuée, ce qui permet de déterminer l'importance de leur impact et leur répartition dans les différents sous-groupes et notamment, le cas échéant, les diverses ethnies qui constituent cette population d'individus.

Les données de génotypage sont analysées pour estimer les
5 fréquences de distributions des différents allèles observés dans les populations étudiées. Les calculs de fréquences alléliques peuvent être réalisés à l'aide de logiciels tels SAS-suite® (SAS) ou SPLUS® (MathSoft). La comparaison des distributions alléliques du polymorphisme de type SNP de l'invention au travers des différentes ethnies de la population d'individus peut être réalisée au moyen
10 des logiciels ARLEQUIN® et SAS-suite®.

La présente invention a également pour objet l'utilisation d'un polynucléotide conforme à l'invention pour la recherche d'une variation dans la séquence nucléotidique du gène IFN α -21 chez un individu.

15 Vecteur d'expression et cellule hôte

La présente invention a aussi pour objet un vecteur recombinant comprenant au moins un polynucléotide conforme à l'invention.

De nombreux systèmes d'expression peuvent être utilisés, comme, par exemple, les chromosomes, les épisomes, les virus dérivés. Plus
20 particulièrement, les vecteurs recombinants utilisés peuvent être dérivés de plasmides bactériens, de transposons, d'épisome de levure, d'éléments d'insertion, d'éléments chromosomiques de levures, de virus tels que les baculovirus, les papillonna virus comme SV40, les vaccinia virus, les adénovirus, les fox pox virus, les pseudorabies virus, les rétrovirus.

25 Ces vecteurs recombinants peuvent également être des dérivés de cosmides ou de phagemides. La séquence nucléotidique peut être insérée dans le vecteur recombinant d'expression par les méthodes bien connues de l'homme du métier, telles que, par exemple, celles qui sont décrites dans MOLECULAR CLONING, A LABORATORY MANUAL (supra) Sambrook et al..

30 Le vecteur recombinant peut comprendre des séquences nucléotidiques de contrôle de la régulation de l'expression du polynucléotide ainsi que des séquences nucléotidiques permettant l'expression et la



transcription d'un polynucléotide de l'invention et la traduction d'un polypeptide de l'invention, ces séquences étant choisies en fonction des cellules hôtes mises en œuvre.

Ainsi, par exemple, un signal de sécrétion approprié peut être
5 intégré dans le vecteur recombinant pour que le polypeptide, codé par le polynucléotide de l'invention, soit dirigé vers la lumière du réticulum endoplasmique, vers l'espace périplasmique, sur la membrane ou vers l'environnement extracellulaire.

La présente invention a aussi pour objet une cellule hôte
10 comprenant un vecteur recombinant conforme à l'invention.

L'introduction du vecteur recombinant dans une cellule hôte peut être effectuée selon les méthodes bien connues de l'homme du métier telles que celles décrites dans BASIC METHODS IN MOLECULAR BIOLOGY, Davis et al., 1986 et MOLECULAR CLONING : A LABORATORY MANUAL, 2nd Ed.,
15 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989, telles que la transfection par calcium phosphate, la transfection par DEAE dextran, la transvection, la microinjection, la transfection par lipides cationiques, l'électroporation, la transduction ou l'infection.

Les cellules hôtes peuvent être, par exemple, des cellules
20 bactériennes telles que les cellules de streptocoque, de staphylocoque, d'*E. coli* ou de *Bacillus subtilis*, des cellules de champignons telles que les cellules de levure et les cellules d'*Aspergillus*, de *Streptomyces*, des cellules d'insectes telles que les cellules de *Drosophila* S2 et de *Spodoptera* Sf9, des cellules animales, telles que les cellules CHO, COS, HeLa, C127, BHK, HEK 293 et des
25 cellules humaines du sujet à traiter ou encore des cellules végétales.

Les cellules hôtes peuvent être utilisées, par exemple, pour exprimer un polypeptide de l'invention ou en tant que produit actif dans des compositions pharmaceutiques, comme on le verra ci-après.

30 Polypeptide

La présente invention a aussi pour objet un polypeptide isolé comprenant une séquence d'acides aminés ayant au moins 80 % d'identité,



préférentiellement au moins 90 % d'identité, plus préférentiellement au moins 95 % d'identité et encore plus préférentiellement au moins 99 % d'identité avec:

- a) la séquence d'acides aminés ID SEQ N° 2, ou avec
- b) la séquence d'acides aminés comportant les acides aminés compris entre 24
5 et 189 de la séquence d'acides aminés ID SEQ N° 2,
sous réserve que chacune des séquences d'acides aminés sous a) et b)
comporte au moins l'un des polymorphismes de type SNP suivants :
 - Q102K,
 - Q114H,
 - 10 - V127D,
 - C162stop,
 - K179E.

Le polypeptide de l'invention peut également comprendre :

- a) la séquence d'acides aminés ID SEQ N° 2, ou
- 15 b) la séquence d'acides aminés comportant les acides aminés compris entre 24
et 189 de la séquence d'acides aminés ID SEQ N° 2,
sous réserve que chacune des séquences d'acides aminés sous a) et b)
comporte au moins l'un des polymorphismes de type SNP suivants :
 - Q102K,
 - 20 - Q114H,
 - V127D,
 - C162stop,
 - K179E.

Le polypeptide de l'invention peut tout particulièrement consister

- 25 en :
 - a) la séquence d'acides aminés ID SEQ N° 2, ou
 - b) la séquence d'acides aminés comportant les acides aminés compris entre 24
et 189 de la séquence d'acides aminés ID SEQ N° 2,
sous réserve que chacune des séquences d'acides aminés sous a) et b)
30 comporte au moins l'un des polymorphismes de type SNP suivants :
 - Q102K,
 - Q114H,



- V127D,
- C162stop,
- K179E.

Conformément à la présente invention, un polypeptide conforme à
5 l'invention ne peut comporter simultanément les polymorphismes de type SNP
C162stop et K179E.

Préférentiellement, un polypeptide conforme à l'invention
comporte un seul polymorphisme de type SNP choisi dans le groupe constitué
par :

- 10 - Q102K,
- Q114H,
- V127D,
- C162stop, et
- K179E.

15 La présente invention a également pour objet un procédé de
préparation d'un polypeptide ci-dessus décrit, dans lequel une cellule hôte
définie précédemment est cultivée et ledit polypeptide est isolé du milieu de
culture.

Le polypeptide peut être purifié à partir des cellules hôtes, selon
20 les méthodes bien connues de l'homme du métier telles que la précipitation à
l'aide d'agents chaotropiques comme les sels, en particulier le sulfate
d'ammonium, l'éthanol l'acétone ou l'acide trichloroacétique, l'extraction à
l'acide; la chromatographie échangeuse d'ions; la chromatographie par
phosphocellulose; la chromatographie par interaction hydrophobe; la
25 chromatographie d'affinité; la chromatographie hydroxylapatite ou les
chromatographies d'exclusion.

On entend par "milieu de culture", le milieu dans lequel on purifie
le polypeptide de l'invention. Ce milieu peut être constitué par le milieu
extracellulaire et/ou le lysat cellulaire. Des techniques bien connues de
30 l'homme du métier permettent également à ce dernier de redonner la
conformation active au polypeptide, si la conformation dudit polypeptide a été
altérée lors de l'isolation ou de la purification.

Anticorps

La présente invention a aussi pour objet un procédé d'obtention d'un anticorps immunospécifique.

5 On entend par "anticorps", les anticorps monoclonaux, polyclonaux, chimériques, simple chaîne, humanisés ainsi que les fragments Fab, incluant les produits d'un Fab ou d'une banque d'expression d'immunoglobulines.

Un anticorps immunospécifique peut être obtenu par immunisation
10 d'un animal avec un polypeptide conforme à l'invention.

L'invention concerne aussi un anticorps immunospécifique pour un polypeptide conforme à l'invention, tel que défini précédemment.

Un polypeptide selon l'invention, un de ses fragments, un analogue, un de ses variants ou une cellule exprimant ce polypeptide peuvent
15 aussi être utilisés pour produire des anticorps immunospécifiques.

Le terme "immunospécifique" signifie que l'anticorps possède une meilleure affinité pour le polypeptide de l'invention que pour d'autres polypeptides connus de l'art antérieur.

Les anticorps immunospécifiques peuvent être obtenus par
20 administration d'un polypeptide de l'invention, d'un de ses fragments, d'un analogue ou d'un fragment épitopique ou d'une cellule exprimant ce polynucléotide chez un mammifère, de préférence non humain, selon les méthodes bien connues de l'homme du métier.

Pour la préparation d'anticorps monoclonaux, on peut utiliser des
25 méthodes usuelles de production d'anticorps, à partir de lignées cellulaires, telles que la technique des hybridomes (Kohler et al., Nature (1975) 256 :495-497), la technique des triomes, la technique des hybridomes de cellules B humaines (Kozbor et al., Immunology Today (1983) 4 :72) et la technique des hybridomes EBV (Cole et al., MONOCLONAL ANTIBODIES AND CANCER
30 THERAPY, pp. 77-96, Alan R. Liss, 1985).

Les techniques de production d'anticorps simple-chaîne telles que décrites, par exemple, dans US N° 4,946, 778 peuvent être également utilisées.



Des animaux transgéniques comme les souris, par exemple, peuvent être également utilisés pour produire des anticorps humanisés.

Agents interagissant avec le polypeptide de l'invention

5 La présente invention a également pour objet un procédé d'identification d'un agent activateur ou inhibiteur d'un polypeptide conforme à l'invention, comprenant :

- a) la mise en présence de cellules hôtes, telles que définies ci-dessus avec un agent à tester, et
- 10 b) la détermination de l'effet activateur, ou inhibiteur, généré par l'agent à tester.

Un polypeptide conforme à l'invention peut ainsi être employé pour un procédé de criblage de composés qui rentrent en interaction avec celui-ci.

Ces composés peuvent être des agents activateurs (agonistes) ou
15 inhibiteurs (antagonistes) de l'activité intrinsèque d'un polypeptide selon l'invention. Ces composés peuvent également être des ligands ou des substrats d'un polypeptide de l'invention. Voir Coligan et al., Current Protocols in Immunology 1 (2), Chapter 5 (1991).

En général, pour mettre en place un tel procédé, il est d'abord
20 souhaitable de produire des cellules hôtes appropriées qui expriment un polypeptide conforme à l'invention. De telles cellules peuvent être, par exemple, des cellules de mammifères, de levures, d'insectes comme *Drosophila* ou de bactéries comme *E. coli*.

Ces cellules ou des extraits de membrane de ces cellules, sont
25 alors mises en présence des composés à tester.

On peut ainsi observer la capacité de liaison des composés à tester avec le polypeptide de l'invention, mais également l'inhibition ou l'activation de la réponse fonctionnelle.

L'étape b) du procédé ci-dessus peut être mise en œuvre en
30 utilisant un agent à tester marqué directement ou indirectement. Elle peut aussi comprendre un test de compétition, en utilisant un agent marqué ou non et un agent compétiteur marqué.

On peut également déterminer si un agent à tester conduit à la génération d'un signal d'activation ou d'inhibition sur des cellules exprimant le polypeptide de l'invention, en utilisant des moyens de détection appropriés, suivant le signal à détecter.

5 De tels agents activateurs ou inhibiteurs peuvent être des polynucléotides, et dans certains cas des oligonucléotides ou des polypeptides, comme des protéines ou des anticorps, par exemple.

La présente invention a aussi pour objet une méthode pour l'identification d'un agent activé ou inhibé par un polypeptide conforme à
10 l'invention, comprenant :

- a) la mise en présence de cellules hôtes obtenues comme décrit ci-dessus avec un agent à tester, et
- b) la détermination de l'effet activateur ou inhibiteur, généré par ledit polypeptide sur l'agent à tester.

15 Un agent activé ou inhibé par le polypeptide de l'invention est un agent qui répond, respectivement, par une activation ou une inhibition en présence de ce polypeptide. Les agents activés ou inhibés, directement ou indirectement, par le polypeptide de l'invention peuvent consister en des polypeptides comme, par exemple, des récepteurs membranaires ou
20 nucléaires, des kinases et plus préférentiellement des tyrosines kinases, des facteurs de transcription ou des polynucléotides.

Détection de maladies

La présente invention a aussi pour objet un procédé de détection
25 de l'expression et/ou de l'activité d'un polypeptide conforme à l'invention, comprenant :

- a) la détection de la présence ou de l'absence d'un polynucléotide selon l'invention dans le génome du sujet, et/ou
- b) la détection de la présence, de l'absence et/ou d'une concentration
30 prédéterminée d'un polypeptide selon l'invention chez un sujet.

La détection d'un polynucléotide et d'un polypeptide de l'invention peut ainsi permettre de savoir si un sujet est atteint ou risque d'être atteint ou,

au contraire, présente une résistance partielle au développement d'une maladie, d'une indisposition ou d'un dérèglement tels que définis précédemment, relatifs à l'expression et/ou l'activité d'un polypeptide de l'invention.

5 La concentration "normale" d'un polypeptide peut être prédéterminée par l'homme du métier par des tests ou des essais conventionnels qui lui permettront d'identifier le seuil au-dessus ou en dessous duquel apparaît la sensibilité ou, au contraire, la résistance à la maladie, l'indisposition ou le dérèglement évoqué ci-dessus.

10 Le polynucléotide à tester peut être obtenu à partir d'échantillons biologiques du sujet à étudier, tels que des cellules, du sang, de l'urine, de la salive, ou à partir d'une biopsie ou d'une autopsie du sujet à étudier. L'ADN génomique peut être utilisé directement pour la détection ou après à une amplification par PCR, par exemple. L'ARN ou l'ADNc peuvent également être
15 utilisés de façon similaire.

Il est ensuite possible de comparer la séquence nucléotidique d'un polynucléotide conforme à l'invention avec la séquence nucléotidique détectée dans le génome du sujet.

La comparaison des séquences nucléotidiques peut être effectuée
20 par séquençage par des méthodes d'hybridation de l'ADN, par différence de mobilité des fragments d'ADN sur gel d'électrophorèse avec ou sans agents dénaturants ou par différence de températures de fusion. Voir Myers et al., Science (1985) 230 :1242. De telles modifications dans la structure de la séquence nucléotidique en un point précis peuvent également être révélées par
25 des essais de protection aux nucléases, telles que l'ARNase et la nucléase S1 ou encore par des agents chimiques de clivage. Voir Cotton et al., Proc. Nat. Acad. Sci. USA (1985) 85 :4397-4401. Des sondes oligonucléotidiques comprenant un fragment d'un polynucléotide de l'invention peuvent également être utilisées pour conduire le criblage.

30 De nombreuses méthodes bien connues de l'homme du métier peuvent être utilisées pour déterminer l'expression d'un polynucléotide de

l'invention et pour identifier la variabilité génétique de ce polynucléotide. Voir Chee et al., Science (1996), Vol 274, pp 610-613.

La présente invention a également pour objet l'utilisation d'un polynucléotide conforme à l'invention pour effectuer un diagnostic génétique d'une maladie ou d'une résistance à une maladie liée à la présence, chez un ou plusieurs individus de la population humaine, de l'allèle mutant codé par le polymorphisme de type SNP fonctionnel conforme à l'invention.

Ces maladies peuvent être des dérèglements et/ou des maladies humaines, tels que les différents cancers comme, par exemple, les carcinomes, les mélanomes, les lymphomes, les myélomes, les tumeurs cancéreuses ou non, les leucémies et les cancers du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que celles qui ne sont pas liées au système immunitaire comme, par exemple, l'obésité, les maladies infectieuses en particulier les infections virales comme les hépatites B et C et le SIDA, les pneumonies, les colites ulcéraives, les maladies du système nerveux central comme, par exemple, la schizophrénie et la dépression, le rejet de greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie en particulier chez les patients dialysés, les allergies, l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les maladies et les désordres auto-immuns, les verrues génitales ou vénériennes, les désordres gastro-intestinaux ou encore les désordres liés aux traitements par chimiothérapie.

Préférentiellement, ces maladies peuvent être des hépatites chroniques B et C, des leucémies telles que la leucémie à trichloroleucocytes et la leucémie myéloïde chronique, des myélomes multiples, des lymphomes folliculaires, des tumeurs carcinoïdes, des mélanomes malins, des carcinomes rénaux métastasés, la maladie d'Alzheimer, la maladie de Parkinson ainsi que des tumeurs qui apparaissent à la suite d'un déficit immunitaire, tel que le sarcome de Kaposi dans le cas du SIDA.

De même, la présence, l'absence et/ou la concentration de polypeptide selon l'invention peuvent aider au diagnostic d'une maladie, une indisposition ou un dérèglement ou, au contraire, la résistance à une maladie,

une indisposition ou un dérèglement chez un sujet en prélevant un échantillon dérivé de ce sujet.

L'augmentation ou la diminution de l'expression du polypeptide peut être mesurée en quantifiant le niveau d'ARN codant pour ce polypeptide, 5 suivant les méthodes bien connues de l'homme du métier, par exemple, par PCR, RT-PCR, protection à l'ARNase, Northern blot, et autres méthodes d'hybridation.

Il est également possible de déterminer la concentration en polypeptides de l'invention présents dans un échantillon biologique du sujet par 10 des méthodes bien connues, par exemple, par radioimmunoessai, tests de liaisons compétitives, Western blot et tests ELISA.

Médicaments et traitements des maladies

Les polypeptides de l'invention possèdent de très intéressantes 15 propriétés pharmacologiques puisqu'ils ont une structure tridimensionnelle et une fonctionnalité différente par rapport à l'IFN α -21 naturel.

Ces propriétés justifient l'utilisation d'un polypeptide conforme à l'invention pour le traitement thérapeutique du corps humain, c'est-à-dire à titre de médicament.

20 C'est pourquoi la présente invention a pour objet un médicament renfermant, à titre de principe actif, un polypeptide conforme à l'invention.

L'invention concerne encore l'utilisation d'un polypeptide conforme à l'invention, pour la fabrication d'un médicament destinée à la prévention ou au traitement de différents dérèglements et/ou maladies humaines, tels que les 25 différents cancers comme, par exemple, les carcinomes, les mélanomes, les lymphomes, les myélomes, les tumeurs cancéreuses ou non, les leucémies et les cancers du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que celles qui ne sont pas liées au système immunitaire comme, par exemple, l'obésité, les maladies 30 infectieuses en particulier les infections virales comme les hépatites B et C et le SIDA, les pneumonies, les colites ulcératives, les maladies du système nerveux central comme, par exemple, la schizophrénie et la dépression, le rejet de

greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie en particulier chez les patients dialysés, les allergies, l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les maladies et les désordres auto-immuns, les verrues génitales ou
5 vénériennes, les désordres gastro-intestinaux ou encore les désordres liés aux traitements par chimiothérapie.

Préférentiellement, le polypeptide conforme à l'invention peut être utilisé pour la fabrication d'un médicament destinée à la prévention ou au traitement de différents dérèglements et/ou maladies humaines, tels que les
10 hépatites chroniques B et C, les leucémies telles que la leucémie à trichoroleucocytes et la leucémie myéloïde chronique, les myélomes multiples, les lymphomes folliculaires, les tumeurs carcinoïdes, les mélanomes malins, les carcinomes rénaux métastasés, la maladie d'Alzheimer, la maladie de Parkinson ainsi que les tumeurs qui apparaissent à la suite d'un déficit
15 immunitaire, tel que le sarcome de Kaposi dans le cas du SIDA.

Certains des composés permettant d'obtenir le polypeptide conforme à l'invention ainsi que les composés obtenus ou identifiés par ou à partir de ce polypeptide peuvent également être utilisés pour le traitement thérapeutique du corps humain, c'est-à-dire à titre de médicament.

20 C'est pourquoi la présente invention a aussi pour objet un médicament contenant, à titre de principe actif un polynucléotide conforme à l'invention, un vecteur recombinant défini précédemment, une cellule hôte définie précédemment, un anticorps défini précédemment et/ou un agent activateur et/ou inhibiteur d'un polypeptide conforme à l'invention.

25 L'invention concerne encore l'utilisation d'un polynucléotide conforme à l'invention, d'un vecteur recombinant défini précédemment, d'une cellule hôte définie précédemment, d'un anticorps défini précédemment et/ou un agent activateur et/ou inhibiteur d'un polypeptide conforme à l'invention, pour la fabrication d'un médicament destinée à la prévention ou au traitement de
30 différents dérèglements et/ou maladies humaines, tels que les différents cancers comme, par exemple, les carcinomes, les mélanomes, les lymphomes, les myélomes, les tumeurs cancéreuses ou non, les leucémies et les cancers

du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que celles qui ne sont pas liées au système immunitaire comme, par exemple, l'obésité, les maladies infectieuses en particulier les infections virales comme les hépatites B et C et le SIDA, les
5 pneumonies, les colites ulcéraives, les maladies du système nerveux central comme, par exemple, la schizophrénie et la dépression, le rejet de greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie en particulier chez les patients dialysés, les allergies, l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les
10 maladies et les désordres auto-immuns, les verrues génitales ou vénériennes, les désordres gastro-intestinaux ou encore les désordres liés aux traitements par chimiothérapie.

Préférentiellement, un polynucléotide conforme à l'invention, un vecteur recombinant défini précédemment, une cellule hôte définie
15 précédemment, un anticorps défini précédemment et/ou un agent activateur et/ou inhibiteur d'un polypeptide conforme à l'invention peuvent être utilisés pour la fabrication d'un médicament destinée à la prévention ou au traitement de différents dérèglements et/ou maladies humaines, tels que les hépatites chroniques B et C, les leucémies telles que la leucémie à trichloroleucocytes et
20 la leucémie myéloïde chronique, les myélomes multiples, les lymphomes folliculaires, les tumeurs carcinoïdes, les mélanomes malins, les carcinomes rénaux métastasés, la maladie d'Alzheimer, la maladie de Parkinson ainsi que les tumeurs qui apparaissent à la suite d'un déficit immunitaire, tel que le sarcome de Kaposi dans le cas du SIDA.

25 Le dosage d'un polypeptide et des autres composés de l'invention, utiles en tant que principe actif, dépend du choix du composé, du mode d'administration, de la nature de la formulation, de la nature du sujet et du jugement du médecin.

Lorsqu'il est utilisé comme principe actif, un polypeptide conforme
30 à l'invention est généralement administré à des doses comprises entre 1 et 100 µg/kg du sujet.

L'invention a aussi pour objet une composition pharmaceutique qui renferment au moins un composé précité, tel qu'un polypeptide conforme à l'invention, un polynucléotide conforme à l'invention, un vecteur recombinant défini précédemment, une cellule hôte définie précédemment, un anticorps
5 défini précédemment et/ou un agent activateur et/ou inhibiteur d'un polypeptide conforme à l'invention, à titre de principe actif, ainsi qu'un excipient pharmaceutiquement acceptable.

Dans ces compositions pharmaceutiques, le principe actif est avantageusement présent à des doses physiologiquement efficaces.

10 Ces compositions pharmaceutiques peuvent être, par exemple, solides ou liquides et se présenter sous les formes pharmaceutiques couramment utilisées en médecine humaine, comme par exemple les comprimés simples ou dragéifiés, les gélules, les granulés, les caramels, les suppositoires et de préférence les préparations injectables et les poudres pour
15 injectables. Ces formes pharmaceutiques peuvent être préparées selon les méthodes usuelles.

Le ou les principes actifs peuvent y être incorporés à des excipients habituellement employés dans ces compositions pharmaceutiques, tels que le talc, la gomme arabique, le lactose, l'amidon, le dextrose, le glycérol,
20 l'éthanol, le stéarate de magnésium, le beurre de cacao, les véhicules aqueux ou non, les corps gras d'origine animale ou végétale, les dérivés paraffiniques, les glycols, les divers agents mouillants, dispersants ou émulsifiants, les conservateurs.

Le ou les principes actifs conforme à l'invention peuvent être
25 employés seul ou en combinaison avec d'autres composés, tels que des composés thérapeutiques tels que d'autres IFN α , voire d'autres cytokines comme les interleukines, par exemple.

Les différentes formulations des compositions pharmaceutiques sont adaptées suivant le mode d'administration.

30 Les compositions pharmaceutiques peuvent être administrées par les différentes voies d'administration connues de l'homme du métier.

L'invention a également pour objet une composition de diagnostic qui renferment au moins un composé précité, tel qu'un polypeptide conforme à l'invention, un polynucléotide conforme à l'invention, un vecteur recombinant défini précédemment, une cellule hôte définie précédemment, un anticorps
5 défini précédemment et/ou un agent activateur et/ou inhibiteur d'un polypeptide conforme à l'invention, à titre de principe actif, ainsi qu'un excipient approprié pharmaceutiquement acceptable.

Les excipients appropriés utilisés dans la composition de diagnostic sont généralement des tampons et des conservateurs.

10 La présente invention a également pour objet l'utilisation :

- a) d'une quantité thérapeutiquement efficace d'un agent activateur défini précédemment et/ou
- b) d'une quantité thérapeutiquement efficace d'un polypeptide selon l'invention, et/ou
- 15 c) d'un polynucléotide selon l'invention, et/ou
- d) d'une cellule hôte du sujet à traiter, définie précédemment, pour préparer un médicament destiné à augmenter l'expression ou l'activité chez un sujet, d'un polypeptide conforme à l'invention.

Ainsi, pour traiter un sujet qui a besoin d'une augmentation de
20 l'expression ou de l'activité d'un polypeptide de l'invention, plusieurs méthodes sont possibles.

Il est possible d'administrer au sujet une quantité thérapeutiquement efficace d'un polypeptide de l'invention et/ou d'un agent activateur et/ou activé tels que définis précédemment, éventuellement en
25 combinaison, avec un excipient pharmaceutiquement acceptable.

Il est également possible d'augmenter la production endogène d'un polypeptide de l'invention par administration au sujet d'un polynucléotide selon l'invention. Par exemple, ce polynucléotide peut être inséré dans un vecteur rétroviral d'expression. Un tel vecteur peut être isolé à partir de cellules
30 ayant été infectées par un vecteur de plasmide rétroviral contenant de l'ARN codant pour le polypeptide de l'invention, de telle façon pour que les cellules transduites produisent des particules virales infectieuses contenant le gène

d'intérêt. Voir Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, Chapter 20, in Human Molecular Genetics, Strachan and Read, BIOS Scientific Publishers Ltd (1996).

5 Il est également possible d'administrer au sujet des cellules hôtes lui appartenant, ces cellules hôtes ayant été prélevées et modifiées, au préalable, de façon à exprimer le polypeptide de l'invention, comme décrit précédemment.

La présente invention concerne également l'utilisation :

- 10 a) d'une quantité thérapeutiquement efficace d'un agent inhibiteur défini précédemment, et/ou
- b) d'une quantité thérapeutiquement efficace d'un anticorps immunospécifique défini précédemment, et/ou
- c) d'un polynucléotide permettant d'inhiber l'expression d'un polynucléotide conforme à l'invention,
- 15 pour préparer un médicament destiné à diminuer l'expression ou l'activité, chez un sujet, d'un polypeptide conforme à l'invention.

Ainsi, il est possible d'administrer au sujet une quantité thérapeutiquement efficace d'un agent inhibiteur et/ou d'un anticorps tels que définis précédemment, éventuellement en combinaison, avec un excipient

20 pharmaceutiquement acceptable.

Il est également possible de diminuer la production endogène d'un polypeptide de l'invention par administration au sujet d'un polynucléotide complémentaire conforme à l'invention permettant d'inhiber l'expression d'un polynucléotide de l'invention.

25

PARTIE EXPERIMENTALE

Exemple 1 : Modélisation d'une protéine codée par un polynucléotide de séquence nucléotidique comportant un polymorphisme de type SNP conforme à l'invention et de la protéine codée par la séquence nucléotidique du gène

30 sauvage de référence

Dans une première étape la structure tridimensionnelle de IFN α -21 a été construite à partir de celle de IFN α -2 dont la structure est disponible

dans la base de données PDB (code 1ITF) et ce en utilisant le logiciel Modeler (MSI, San Diego, CA).

Le fragment polypeptidique mature a ensuite été modifié de façon à reproduire la mutation Q102K, Q114H, V127D, C162stop ou K179E.

5 Un millier d'étapes de minimisations moléculaires ont été conduites sur les fragments mutés en utilisant les programmes AMBER et DISCOVER (MSI : Molecular Simulations Inc.).

Deux suites de calculs de dynamiques moléculaires ont ensuite été effectuées avec le même programme et les mêmes champs de forces.

10 Dans chaque cas, 50000 étapes ont été calculées à 300°K, terminées par 300 étapes d'équilibration.

Le résultat de cette modélisation est visualisé sur les Figures 1, 2, 3, 4, 5, 6, 7, 8 et 9.

15 Exemple 2 : Génotypage d'un polymorphisme de type SNP conforme à l'invention dans une population d'individus

Le génotypage de SNPs est basé sur le principe du miniséquençage dont le produit est détecté par une lecture de fluorescence polarisée. La technique consiste en un miniséquençage fluorescent
20 (Technologie FP-TDI ou Fluorescence Polarization Template-direct Dye-terminator Incorporation).

Le miniséquençage consiste à allonger un oligonucléotide amorce, placé juste en amont du site polymorphe, par des didéoxynucléotides fluoromarkés à l'aide d'une enzyme polymérase. Le résultat de cet
25 allongement est directement analysé par une lecture de fluorescence polarisée.

A) Protocole

Le miniséquençage est réalisé sur un produit obtenu après amplification par PCR à partir de l'ADN génomique de chaque individu d'une
30 population d'individus (définie ci-après), d'un fragment de séquence nucléotidique de l'IFN α -21.

Ce produit PCR est choisi pour couvrir la région génique



contenant le polymorphisme de type SNP de l'invention. Ensuite, on élimine les amorces de PCR et les dNTPs non incorporés avant de réaliser le miniséquençage. Toutes ces étapes, ainsi que la lecture, sont réalisées dans la même plaque.

5 Le génotypage requiert donc 5 étapes :

- 1) Amplification par PCR
- 2) Purification du produit de PCR par digestion enzymatique
- 3) Elongation de l'oligonucléotide amorce
- 4) Lecture

10 5) Interprétation de la lecture

Les étapes de génotypage 1 et 2 sont les mêmes pour les polymorphismes de type SNP : c973a, g1011c, t1049a, t1155a et a1204g. Les étapes 3, 4 et 5 sont spécifiques pour chaque polymorphisme de type SNP de l'invention.

15

Etape 1)

L'amplification PCR de la séquence nucléotidique du gène IFN α -21 qui couvre la région génique contenant le polymorphisme de type SNP fonctionnel conforme à l'invention est effectuée pour chaque individu de la population d'individus.

20

La population d'individus est composée d'ADNs génomiques fournis par l'Institut Coriell aux Etats-Unis.

Les 239 individus choisis aléatoirement se répartissent comme suit :

25



| POPULATION | DESCRIPTION | NOMBRE D'INDIVIDUS |
|------------|-------------------------|--------------------|
| 1 | Afro Américain | 50 |
| 2 | Amérindien du Sud Ouest | 5 |
| 3 | Sud Américain (Andes) | 10 |
| 4 | Caribéen | 10 |
| 5 | Caucasien | 50 |
| 6 | Chinois | 10 |
| 7 | Grec | 8 |
| 8 | Ibérien | 10 |
| 9 | Italien | 10 |
| 10 | Japonais | 10 |
| 11 | Mexicain | 10 |
| 12 | Moyen-Orient | 20 |
| 13 | Individus du Pacifique | 7 |
| 14 | Indo-Pakistanaï | 9 |
| 15 | Sud Américain | 10 |
| 16 | Asie du Sud | 10 |

L'amplification PCR est réalisée à partir d'amorces que l'homme du métier peut facilement déterminer à l'aide de la séquence nucléotidique ID SEQ N° 1.

5 Les amorces sont les suivantes : ID SEQ N° 3 et ID SEQ N° 4.

Ces séquences nucléotidiques permettent d'amplifier un fragment d'une longueur de 696 nucléotides, du nucléotide 620 au nucléotide 1315 dans la séquence nucléotidique ID SEQ N° 1

10 Les amplifiats PCR serviront de matrice pour la réaction de miniséquençage.

Le volume réactionnel est de 5 µl par échantillon comme décrit dans le tableau suivant :

| Fournisseur | Référence | Réactif | Conc. initiale | Vol. par tube (µl) | Conc. finale |
|------------------|----------------|-------------------------------|----------------|---------------------|--------------------|
| Life Technologie | Livré avec Taq | Tampon (X) | 10 | 0,5 | 1 |
| Life Technologie | Livré avec Taq | MgSO ₄ (mM) | 50 | 0,2 | 2 |
| AP Biotech | 27-2035-03 | dNTPs (mM) | 10 | 0,1 | 0,2 |
| | Sur demande | Amorce F (µM) ID SEQ N° 4 | 10 | 0,1 | 0,2 |
| | Sur demande | AmorceR (µM) ID SEQ N° 5 | 10 | 0,1 | 0,2 |
| Life Technologie | 11304-029 | Taq platinum | 5U/ µl | 0,02 | 0,1 U/ réaction |
| | | H ₂ O | Qsp 5 µl | 1,98 | |
| | | ADN | 2,5 ng/ µl | 2 | 5 ng/ réaction |
| | | Volume total | | 5 µl | |

Ces réactifs sont distribués dans une plaque PCR noire à 384 puits fournie par ABGene (ref :TF-0384-k). Une fois remplie, la plaque est scellée, centrifugée puis placée dans un thermocycleur pour plaque 384 (Tetrad de MJ Research) et subit l'incubation suivante : Cycles de PCR : 1 min à 94° C, suivi de 36 cycles composés de 3 étapes (15 sec. à 94° C, 30 sec. à 56° C, 1 min. à 68° C)

Etape 2)

La PCR est ensuite purifiée à l'aide de deux enzymes que sont la phosphatase alcaline de crevette (ou Shrimp Alkaline Phosphatase SAP) et l'exonucléase I (Exo I). La première de ces enzymes permet la déphosphorylation des dNTPs non incorporés au cours de la PCR, tandis que la seconde élimine les résidus simple brin d'ADN et donc les amorces non utilisées au cours de la PCR. Cette digestion se fait par addition dans la plaque de PCR d'un mélange réactionnel de 5 µl par échantillon préparé comme décrit dans le tableau suivant :



| Fournisseur | Référence | Réactif | Conc. initiale | Vol. par tube (µl) | Conc. finale |
|-------------|-----------------|------------------|----------------|---------------------|---------------|
| AP Biotech | E70092X | SAP | 1 U/ µl | 0,5 | 0,5/ réaction |
| AP Biotech | 070073Z | Exo I | 10 U/ µl | 0,1 | 1/ réaction |
| AP Biotech | Fourni avec SAP | Tampon SAP (X) | 10 | 0,5 | 1 |
| | | H ₂ O | Qsp 5 µl | 3,9 | |
| | | PCR | | 5 µl | |
| | | Vol total | | 10 µl | |

Une fois remplie, la plaque est scellée, centrifugée puis placée dans un thermocycleur pour plaque 384 puits (Tetrad de MJ Research) et subit l'incubation suivante : Digestion SAP-EXO : 45 min à 37° C, 15min à 80° C.

5

Etape 3)

L'étape d'élongation ou de miniséquençage est ensuite réalisée sur ce produit de PCR digéré, par addition d'un mélange réactionnel de 5 µL par échantillon préparé.

10

Le miniséquençage est spécifique à chaque polymorphisme de type SNP : c973a, g1011c, t1049a, t1155a et a1204g.

A) Polymorphisme de type SNP c973a

L'étape d'élongation ou de miniséquençage est ensuite réalisée
15 comme indiqué dans le tableau suivant :

| Fournisseur | Référence | Réactif | Conc. initiale | Vol. par tube (µl) | Conc. finale |
|--------------------|---------------------------|--|------------------|--------------------|--------------------|
| Propre préparation | | Tampon Elongation ¹ (X) | 5 | 1 | 1 |
| Life Technologies | Sur demande | Amorce Miniseq (µM) A ou B | 10 | 0,5 | 1 |
| AP Biotech | 27-2051 (61,71,81)-01 | ddNTPs ² (µM) 2 non marqués | 2,5 de chaque | 025 | 0,125 de chaque |
| NEN | Nel 472/5 et Nel 492/5 | ddNTPs ² (µM) 2 marqués Tamra et R110 | 2,5 de chaque | 0,25 | 0,125 de chaque |
| AP Biotech | E79000Z | Thermo-sequenase | 3,2 U/ µl | 0,125 | 0,4 U/ réaction |
| | | H2O | Qsp 5 µl | 3,125 | |
| | | PCR digérée | | 10 µl | |
| | | Vol total | | 15 µl | |

¹ Le tampon élongation : Le tampon élongation 5X est composé de Tris-HCl pH 9 à 250 mM, de KCl à 250 mM, de NaCl à 25 mM, de MgCl₂ à 10 mM et de glycérol à 40 %.

- ² ddNTPs : Pour les ddNTPs, un mélange des 4 bases est réalisé en fonction du polymorphisme étudié. Seulement les 2 bases d'intérêts (G/T) composant le polymorphisme de type SNP fonctionnel portent un marquage, soit en Tamra, soit en R110. Le mélange de ddNTPs est composé de :
- 2,5 µM de ddATP non marqué,
 - 2,5 µM de ddCTP non marqué,
 - 10 - 2,5 µM de ddTTP (1,825 µM de ddTTP non marqué et 0,625 µM de ddTTP marqué au Tamra),
 - 2,5 µM de ddGTP (1,825 µM de ddGTP non marqué et 0,625 µM de ddGTP marqué au R110).

Une fois remplie, la plaque est scellée, centrifugée puis placée dans un thermocycleur pour plaque 384 puits (Tetrad de MJ Research) et subit l'incubation suivante : Cycles d'élongation : 1 min. à 93° C, suivi de 35 cycles composés de 2 étapes (10 sec. à 93° C, 30 sec. à 55° C).

Après la dernière étape dans le thermocycleur, la plaque est directement placée sur un lecteur de fluorescence polarisée de type Analyst® HT de LJI Biosystems Inc.. La plaque est lue à l'aide du logiciel Criterion Host® en utilisant deux méthodes. La première permet de lire la base marquée en Tamra en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 550-10 nm, émission 580-10 nm) et la seconde



permet de lire la base marquée en R110 en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 490-10 nm, émission 520-10 nm). Dans les deux cas, un miroir double dichroïque (R110/Tamra) est utilisé et les autres paramètres de lecture sont :

5 Z-height : 1,5 mm

Attenuator : out

Temps d'intégration : 100,000 µsec.

Raw data units : counts/sec

Switch polarization : by well

10 Plate settling time : 0 msec

PMT setup : Smart Read (+), sensitivity 2

Dynamic polarizer : emission

Static polarizer : S

Un fichier résultat est alors obtenu contenant les valeurs calculées de mP pour le filtre Tamra et celle pour le filtre R110. Ces valeurs de mP sont calculées à partir des valeurs d'intensité obtenues sur le plan parallèle(//) et sur le plan perpendiculaire (⊥) d'après la formule suivante :

$$mP = 1000(// - g\perp)/(// + g\perp).$$

Dans ce calcul, la valeur sur le filtre ⊥ est pondérée d'un facteur g.

20 Celui-ci est un paramètre machine qui doit être déterminé préalablement expérimentalement.

Etapes 4) et 5) Interprétation de la lecture et détermination des génotypes.

Les valeurs de mP sont reportées sur un graphe à l'aide du logiciel Excel de Microsoft Inc., et/ou du logiciel Allele Caller® développé par LJJL Biosystems Inc..

En abscisse est indiquée la valeur de mP de la base marquée au Tamra, en ordonnée est indiquée la valeur de mP de la base marquée au R110. Une forte valeur de mP indique que la base marquée avec ce fluorophore est incorporée et, inversement, une faible valeur de mP révèle l'absence d'incorporation de cette base.

On obtient jusqu'à 3 groupes homogènes de séquences



nucléotidiques ayant des génotypes différents, comme indiqué dans la Figure 10.

L'utilisation du logiciel Allele Caller® permet, une fois le repérage des différents groupes réalisé, d'extraire directement le génotype défini pour chaque individu sous forme d'un tableau.

Les séquences des deux amorces de miniséquençage nécessaires pour le génotypage ont été déterminées. Ces amorces sont sélectionnées pour correspondre à une vingtaine voire une trentaine de nucléotides placés juste en amont du site polymorphe. Du fait que le produit de PCR contenant un SNP est un produit d'ADN double brin, le génotypage peut donc se faire soit sur le brin sens soit sur le brin antisens. Les amorces sélectionnées sont fabriquées par Life Technologies Inc.

Les amorces du miniséquençage sont les suivantes :

Amorce sens : (A) : actcatctgctacttgggaa

Amorce antisens : (B) : aaattttctaggaggctct

Le miniséquençage du polymorphisme de type SNP c973a a d'abord été validé sur 16 échantillons, puis génotypé sur l'ensemble de la population d'individus composée de 239 individus et 7 blancs.

Conditions de miniséquençage testées :

Condition N° 1 :

Amorce sens + ddCTP-R110 + ddATP-Tamra

Condition N° 2 :

Amorce sens + ddATP-R110 + ddCTP-Tamra

Condition N° 3 :

Amorce antisens + ddGTP-R110 + ddTTP-Tamra

Condition N° 4 :

Amorce antisens + ddTTP-R110 + ddGTP-Tamra

Ces 4 conditions ont été testées et la condition N° 3 a été retenue pour le génotypage.

B) Résultats

Après la réalisation complète du processus de génotypage, la



détermination des génotypes des individus de la population d'individus pour le SNP fonctionnel étudié ici a été réalisée à l'aide du graphe représenté sur la Figure 10.

5 Ce génotype est en théorie soit homozygote GG, soit hétérozygote GT, soit homozygote TT chez les individus testés. En réalité et comme montré ci-dessous, le génotype homozygote TT n'est pas détecté dans la population d'individus.

10 Les résultats des contrôles, de la répartition des génotypes déterminés dans la population d'individus et le calcul des différentes fréquences alléliques pour ce polymorphisme de type SNP fonctionnel sont présentés dans les tableaux suivants :

| Nombre d'individus | | Nombre de blanc | | Pourcentage de réussite |
|--------------------|-----------|-----------------|---------|-------------------------|
| testés | génotypés | testés | validés | |
| 239 | 237 | 7 | 7 | 99,2 |



| POPULATION | N | % | Fréquence allélique | | | Génotype TT | | Génotype TG | | Génotype GG | |
|-------------------------|-----|------|---------------------|--------|------|-------------|---|-------------|------|-------------|------|
| | | | % | 95% IC | | N | % | N | % | N | % |
| Afro Américain | 50 | 20,9 | 0 | - | - | 0 | 0 | 0 | 0 | 49 | 100 |
| Amérindien du Sud Ouest | 5 | 2,1 | 0 | - | - | 0 | 0 | 0 | 0 | 5 | 100 |
| Sud Américain (Andes) | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Caribéen | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Caucasien | 50 | 20,9 | 3,0 | 0,0 | 6,3 | 0 | 0 | 3 | 6,0 | 47 | 94,0 |
| Chinois | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Grec | 8 | 3,3 | 6,3 | 0,0 | 18,1 | 0 | 0 | 1 | 12,5 | 7 | 87,5 |
| Ibérien | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Italien | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Japonais | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Mexicain | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 9 | 100 |
| Moyen-Orient | 20 | 8,4 | 0 | - | - | 0 | 0 | 0 | 0 | 20 | 100 |
| Individus du Pacifique | 7 | 2,9 | 0 | - | - | 0 | 0 | 0 | 0 | 7 | 100 |
| Indo-Pakistanaï | 9 | 3,8 | 0 | - | - | 0 | 0 | 0 | 0 | 9 | 100 |
| Sud Américain | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Asie du Sud | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Total | 239 | 100 | 0,8 | 0,0 | 1,7 | 0 | 0 | 4 | 1,7 | 233 | 98,3 |



Dans le tableau ci-dessus,

- N représente le nombre d'individus,
- % représente le pourcentage d'individus dans la sous-population spécifique,
- la fréquence allélique représente le pourcentage de l'allèle muté dans la sous-population spécifique,
- 95 % IC représente l'intervalle minimal et maximal de confiance à 95 %.

Il faut préciser que l'allèle G lu en antisens correspond à l'allèle C lu en sens, soit à la présence d'un Q en position 102 de la séquence nucléotidique du gène IFNa21 et donc que l'allèle A lu en antisens correspond à l'allèle T lu en sens correspondant à un K pour cette position dans la séquence de la protéine correspondante.

En examinant ces résultats par population, on constate que les 4 individus hétérozygotes GT sont issus des sous-populations grecque et caucasienne de la population d'individus.

B) Polymorphisme de type SNP g1011c

L'étape d'élongation ou de miniséquençage est ensuite réalisée comme indiqué dans le tableau suivant :

| Fournisseur | Référence | Réactif | Conc. initiale | Vol. par tube (µl) | Conc. finale |
|--------------------|---------------------------|--|------------------|--------------------|--------------------|
| Propre préparation | | Tampon Elongation ¹ (X) | 5 | 1 | 1 |
| Life Technologies | Sur demande | Amorce Miniseq (µM) C ou D | 10 | 0,5 | 1 |
| AP Biotech | 27-2051 (61,71,81)-01 | ddNTPs ² (µM) 2 non marqués | 2,5 de chaque | 025 | 0,125 de chaque |
| NEN | Nel 472/5 et Nel 492/5 | ddNTPs ² (µM) 2 marqués Tamra et R110 | 2,5 de chaque | 0,25 | 0,125 de chaque |
| AP Biotech | E79000Z | Thermo-sequenase | 3,2 U/ µl | 0,125 | 0,4 U/ réaction |
| | | H2O | Qsp 5 µl | 3,125 | |
| | | PCR digérée | | 10 µl | |
| | | Vol total | | 15 µl | |



- 1 Le tampon élongation : Le tampon élongation 5X est composé de Tris-HCl pH 9 à 250 mM, de KCl à 250 mM, de NaCl à 25 mM, de $MgCl_2$ à 10 mM et de glycérol à 40 %.
- 2 ddNTPs : Pour les ddNTPs, un mélange des 4 bases est réalisé en fonction du polymorphisme étudié. Seulement les 2 bases d'intérêts (C/G) composant le polymorphisme de type SNP
- 5 fonctionnel portent un marquage, soit en Tamra, soit en R110. Le mélange de ddNTPs est composé de :
- 2,5 μM de ddATP non marqué,
 - 2,5 μM de ddTTP non marqué,
 - 2,5 μM de ddCTP (1,825 μM de ddCTP non marqué et 0,625 μM de ddCTP marqué au Tamra),
 - 10 - 2,5 μM de ddGTP (1,825 μM de ddGTP non marqué et 0,625 μM de ddGTP marqué au R110).

Une fois remplie, la plaque est scellée, centrifugée puis placée dans un thermocycleur pour plaque 384 puits (Tetrad de MJ Research) et subit l'incubation suivante : Cycles d'élongation : 1 min. à 93° C, suivi de 35 cycles

15 composés de 2 étapes (10 sec. à 93° C, 30 sec. à 55° C).

Après la dernière étape dans le thermocycleur, la plaque est directement placée sur un lecteur de fluorescence polarisée de type Analyst® HT de LJJL Biosystems Inc.. La plaque est lue à l'aide du logiciel Criterion Host® en utilisant deux méthodes. La première permet de lire la base

20 marquée en Tamra en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 550-10 nm, émission 580-10 nm) et la seconde permet de lire la base marquée en R110 en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 490-10 nm, émission 520-10 nm). Dans les deux cas, un miroir double dichroïque (R110/Tamra) est

25 utilisé et les autres paramètres de lecture sont :

Z-height : 1,5 mm

Attenuator : out

Temps d'intégration : 100,000 μsec .

Raw data units : counts/sec

30 Switch polarization : by well

Plate settling time : 0 msec

PMT setup : Smart Read (+), sensitivity 2

Dynamic polarizer : emission

Static polarizer : S



Un fichier résultat est alors obtenu contenant les valeurs calculées de mP pour le filtre Tamra et celle pour le filtre R110. Ces valeurs de mP sont calculées à partir des valeurs d'intensité obtenues sur le plan parallèle(//) et sur le plan perpendiculaire (⊥) d'après la formule suivante :

5
$$mP = 1000(// - g\perp)/(// + g\perp).$$

Dans ce calcul, la valeur sur le filtre ⊥ est pondérée d'un facteur g. Celui-ci est un paramètre machine qui doit être déterminé préalablement expérimentalement.

10 Etapes 4) et 5) Interprétation de la lecture et détermination des génotypes.

Les valeurs de mP sont reportées sur un graphe à l'aide du logiciel Excel de Microsoft Inc., et/ou du logiciel Allele Caller® développé par LJI Biosystems Inc..

15 En abscisse est indiquée la valeur de mP de la base marquée au Tamra, en ordonnée est indiquée la valeur de mP de la base marquée au R110. Une forte valeur de mP indique que la base marquée avec ce fluorophore est incorporée et, inversement, une faible valeur de mP révèle l'absence d'incorporation de cette base.

20 On obtient jusqu'à 3 groupes homogènes de séquences nucléotidiques ayant des génotypes différents, comme indiqué dans la Figure 11.

L'utilisation du logiciel Allele Caller® permet, une fois le repérage des différents groupes réalisé, d'extraire directement le génotype défini pour chaque individu sous forme d'un tableau.

25 Les séquences des deux amorces de miniséquençage nécessaires pour le génotypage ont été déterminées. Ces amorces sont sélectionnées pour correspondre à une vingtaine voire une trentaine de nucléotides placés juste en amont du site polymorphe. Du fait que le produit de PCR contenant un SNP est un produit d'ADN double brin, le génotypage peut
30 donc se faire soit sur le brin sens soit sur le brin antisens. Les amorces sélectionnées sont fabriquées par Life Technologies Inc.

Les amorces du miniséquençage sont les suivantes :



Amorce sens : (C) : tttccactgaacttaacca

Amorce antisens : (D) : gcttcagggtcattcagctg

Le miniséquençage du polymorphisme de type SNP g1011c a d'abord été validé sur 16 échantillons, puis génotypé sur l'ensemble de la population d'individus composée de 239 individus et 7 blancs.

Conditions de miniséquençage testées :

Condition N° 1 :

Amorce sens + ddGTP-R110 + ddCTP-Tamra

Condition N° 2 :

10 Amorce sens + ddCTP-R110 + ddGTP-Tamra

Condition N° 3 :

Amorce antisens + ddGTP-R110 + ddCTP-Tamra

Condition N° 4 :

Amorce antisens + ddCTP-R110 + ddGTP-Tamra

15 Ces 4 conditions ont été testées et la condition N° 3 a été retenue pour le génotypage.

B) Résultats

Après la réalisation complète du processus de génotypage, la détermination des génotypes des individus de la population d'individus pour le SNP fonctionnel étudié ici a été réalisée à l'aide du graphe représenté sur la Figure 11.

Ce génotype est en théorie soit homozygote CC, soit hétérozygote CG, soit homozygote GG chez les individus testés. En réalité et comme montré ci-dessous, le génotype homozygote GG n'est pas détecté dans la population d'individus.

Les résultats des contrôles, de la répartition des génotypes déterminés dans la population d'individus et le calcul des différentes fréquences alléliques pour ce polymorphisme de type SNP fonctionnel sont présentés dans les tableaux suivants :



52

| Nombre d'individus | | Nombre de blanc | | Pourcentage de réussite |
|--------------------|-----------|-----------------|---------|-------------------------|
| testés | génotypés | Testés | validés | |
| 239 | 237 | 7 | 7 | 99,2 |

| POPULATION | N | % | Fréquence allélique | | | Génotype GG | | Génotype GC | | Génotype CC | |
|-------------------------|-----|------|---------------------|--------|-----|-------------|---|-------------|-----|-------------|------|
| | | | % | 95% IC | | N | % | N | % | N | % |
| Afro Américain | 50 | 20,9 | 1,0 | 0,0 | 3,0 | 0 | 0 | 1 | 2,0 | 49 | 98,0 |
| Amérindien du Sud Ouest | 5 | 2,1 | 0 | - | - | 0 | 0 | 0 | 0 | 5 | 100 |
| Sud Américain (Andes) | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Caribéen | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Caucasien | 50 | 20,9 | 0 | - | - | 0 | 0 | 0 | 0 | 50 | 100 |
| Chinois | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Grec | 8 | 3,3 | 0 | - | - | 0 | 0 | 0 | 0 | 8 | 100 |
| Ibérien | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Italien | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Japonais | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Mexicain | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Moyen-Orient | 20 | 8,4 | 0 | - | - | 0 | 0 | 0 | 0 | 20 | 100 |
| Individus du Pacifique | 7 | 2,9 | 0 | - | - | 0 | 0 | 0 | 0 | 7 | 100 |
| Indo-Pakistanaï | 9 | 3,8 | 0 | - | - | 0 | 0 | 0 | 0 | 9 | 100 |
| Sud Américain | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Asie du Sud | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Total | 239 | 100 | 0,2 | 0,0 | 0,6 | 0 | 0 | 1 | 0,4 | 236 | 99,6 |



Dans le tableau ci-dessus,

- N représente le nombre d'individus,
- % représente le pourcentage d'individus dans la sous-population spécifique,
- la fréquence allélique représente le pourcentage de l'allèle muté dans la sous-population spécifique,
- 95 % IC représente l'intervalle minimal et maximal de confiance à 95 %.

Il faut préciser que l'allèle C lu en antisens correspond à l'allèle G lu en sens, soit à la présence d'un Q en position 114 de la séquence nucléotidique du gène IFN α -21 et donc que l'allèle G lu en antisens correspond à l'allèle C lu en sens correspondant à un K pour cette position dans la séquence de la protéine correspondante.

En examinant ces résultats par population, on constate que le seul individu hétérozygote CG est issu de la sous-population afro-américaine de la population d'individus.

C) Polymorphisme de type SNP t1049a

L'étape d'élongation ou de miniséquençage est ensuite réalisée comme indiqué dans le tableau suivant :

| Fournisseur | Référence | Réactif | Conc. initiale | Vol. par tube (μ l) | Conc. finale |
|--------------------|---------------------------|--|------------------|--------------------------|--------------------|
| Propre préparation | | Tampon Elongation ¹ (X) | 5 | 1 | 1 |
| Life Technologies | Sur demande | Amorce Miniseq (μ M) E ou F | 10 | 0,5 | 1 |
| AP Biotech | 27-2051 (61,71,81)-01 | ddNTPs ² (μ M) 2 non marqués | 2,5 de chaque | 0,25 | 0,125 de chaque |
| NEN | Nel 472/5 et Nel 492/5 | ddNTPs ² (μ M) 2 marqués Tamra et R110 | 2,5 de chaque | 0,25 | 0,125 de chaque |
| AP Biotech | E79000Z | Thermo-sequenase | 3,2 U/ μ l | 0,125 | 0,4 U/ réaction |
| | | H ₂ O | Qsp 5 μ l | 3,125 | |
| | | PCR digérée | | 10 μ l | |
| | | Vol total | | 15 μ l | |



- 1 Le tampon élongation : Le tampon élongation 5X est composé de Tris-HCl pH 9 à 250 mM, de KCl à 250 mM, de NaCl à 25 mM, de $MgCl_2$ à 10 mM et de glycérol à 40 %.
- 2 ddNTPs : Pour les ddNTPs, un mélange des 4 bases est réalisé en fonction du polymorphisme étudié. Seulement les 2 bases d'intérêts (A/T) composant le polymorphisme de type SNP
- 5 fonctionnel portent un marquage, soit en Tamra, soit en R110. Le mélange de ddNTPs est composé de :
- 2,5 μM de ddCTP non marqué,
 - 2,5 μM de ddGTP non marqué,
 - 2,5 μM de ddTTP (1,825 μM de ddTTP non marqué et 0,625 μM de ddTTP marqué au Tamra),
 - 10 - 2,5 μM de ddATP (1,825 μM de ddATP non marqué et 0,625 μM de ddATP marqué au R110).

Une fois remplie, la plaque est scellée, centrifugée puis placée dans un thermocycleur pour plaque 384 puits (Tetrad de MJ Research) et subit l'incubation suivante : Cycles d'élongation : 1 min. à 93° C, suivi de 35 cycles

15 composés de 2 étapes (10 sec. à 93° C, 30 sec. à 55° C).

Après la dernière étape dans le thermocycleur, la plaque est directement placée sur un lecteur de fluorescence polarisée de type Analyst® HT de LJJL Biosystems Inc.. La plaque est lue à l'aide du logiciel Criterion Host® en utilisant deux méthodes. La première permet de lire la base

20 marquée en Tamra en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 550-10 nm, émission 580-10 nm) et la seconde permet de lire la base marquée en R110 en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 490-10 nm, émission 520-10 nm). Dans les deux cas, un miroir double dichroïque (R110/Tamra) est

25 utilisé et les autres paramètres de lecture sont :

Z-height : 1,5 mm

Attenuator : out

Temps d'intégration : 100,000 μsec .

Raw data units : counts/sec

30 Switch polarization : by well

Plate settling time : 0 msec

PMT setup : Smart Read (+), sensitivity 2

Dynamic polarizer : emission

Static polarizer : S



Un fichier résultat est alors obtenu contenant les valeurs calculées de mP pour le filtre Tamra et celle pour le filtre R110. Ces valeurs de mP sont calculées à partir des valeurs d'intensité obtenues sur le plan parallèle(//) et sur le plan perpendiculaire (⊥) d'après la formule suivante :

5
$$mP = 1000(// - g\perp)/(// + g\perp).$$

Dans ce calcul, la valeur sur le filtre ⊥ est pondérée d'un facteur g. Celui-ci est un paramètre machine qui doit être déterminé préalablement expérimentalement.

10 Etapas 4) et 5) Interprétation de la lecture et détermination des génotypes.

Les valeurs de mP sont reportées sur un graphe à l'aide du logiciel Excel de Microsoft Inc., et/ou du logiciel Allele Caller® développé par LJI Biosystems Inc..

15 En abscisse est indiquée la valeur de mP de la base marquée au Tamra, en ordonnée est indiquée la valeur de mP de la base marquée au R110. Une forte valeur de mP indique que la base marquée avec ce fluorophore est incorporée et, inversement, une faible valeur de mP révèle l'absence d'incorporation de cette base.

20 On obtient jusqu'à 3 groupes homogènes de séquences nucléotidiques ayant des génotypes différents, comme indiqué dans la Figure 12.

L'utilisation du logiciel Allele Caller® permet, une fois le repérage des différents groupes réalisé, d'extraire directement le génotype défini pour chaque individu sous forme d'un tableau.

25 Les séquences des deux amorces de miniséquençage nécessaires pour le génotypage ont été déterminées. Ces amorces sont sélectionnées pour correspondre à une vingtaine voire une trentaine de nucléotides placés juste en amont du site polymorphe. Du fait que le produit de PCR contenant un SNP est un produit d'ADN double brin, le génotypage peut
30 donc se faire soit sur le brin sens soit sur le brin antisens. Les amorces sélectionnées sont fabriquées par Life Technologies Inc.

Les amorces du miniséquençage sont les suivantes :



Amorce sens : (E) : agcctgcgtgatacaggagg

Amorce antisens : (F) : ggggagtctctccacccca

Le miniséquençage du polymorphisme de type SNP t1049a a d'abord été validé sur 16 échantillons, puis génotypé sur l'ensemble de la population d'individus composée de 239 individus et 7 blancs.

Conditions de miniséquençage testées :

Condition N° 1 :

Amorce sens + ddATP-R110 + ddTTP-Tamra

Condition N° 2 :

10 Amorce sens + ddTTP-R110 + ddATP-Tamra

Condition N° 3 :

Amorce antisens + ddTTP-R110 + ddATP-Tamra

Condition N° 4 :

Amorce antisens + ddATP-R110 + ddTTP-Tamra

15 Ces 4 conditions ont été testées et la condition N° 4 a été retenue pour le génotypage.

B) Résultats

20 Après la réalisation complète du processus de génotypage, la détermination des génotypes des individus de la population d'individus pour le SNP fonctionnel étudié ici a été réalisée à l'aide du graphe représenté sur la Figure 12.

25 Ce génotype est en théorie soit homozygote TT, soit hétérozygote TA, soit homozygote AA chez les individus testés. En réalité et comme montré ci-dessous, le génotype homozygote AA n'est pas détecté dans la population d'individus.

30 Les résultats des contrôles, de la répartition des génotypes déterminés dans la population d'individus et le calcul des différentes fréquences alléliques pour ce polymorphisme de type SNP fonctionnel sont présentés dans les tableaux suivants :



58

| Nombre d'individus | | Nombre de blanc | | Pourcentage de réussite |
|--------------------|-----------|-----------------|---------|-------------------------|
| testés | génotypés | Testés | validés | |
| 239 | 237 | 7 | 7 | 99,2 |



| POPULATION | N | % | Fréquence allélique | | | Génotype AA | | Génotype AG | | Génotype GG | |
|-------------------------|-----|------|---------------------|--------|-----|-------------|-----|-------------|-----|-------------|-------|
| | | | % | 95% IC | | N | % | N | % | N | % |
| Afro Américain | 50 | 20,9 | 1,0 | 0,0 | 3,0 | 0 | 0,0 | 1 | 2,0 | 48 | 98,0 |
| Amérindien du Sud Ouest | 5 | 2,1 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 5 | 100,0 |
| Sud Américain (Andes) | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 9 | 100,0 |
| Caribéen | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 10 | 100,0 |
| Caucasien | 50 | 20,9 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 50 | 100,0 |
| Chinois | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 10 | 100,0 |
| Grec | 8 | 3,3 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 8 | 100,0 |
| Ibérien | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 10 | 100,0 |
| Italien | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 10 | 100,0 |
| Japonais | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 10 | 100,0 |
| Mexicain | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 10 | 100,0 |
| Moyen-Orient | 20 | 8,4 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 20 | 100,0 |
| Individus du Pacifique | 7 | 2,9 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 7 | 100,0 |
| Indo-Pakistanaï | 9 | 3,8 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 9 | 100,0 |
| Sud Américain | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 10 | 100,0 |
| Asie du Sud | 10 | 4,2 | 0,0 | - | - | 0 | 0,0 | 0 | 0,0 | 10 | 100,0 |
| Total | 239 | 100 | 0,2 | 0,0 | 0,6 | 0 | 0,0 | 1 | 0,4 | 236 | 99,6 |



Dans le tableau ci-dessus,

- N représente le nombre d'individus,
- % représente le pourcentage d'individus dans la sous-population spécifique,
- la fréquence allélique représente le pourcentage de l'allèle muté dans la sous-population spécifique,
- 95 % IC représente l'intervalle minimal et maximal de confiance à 95 %.

Il faut préciser que l'allèle A lu en antisens correspond à l'allèle T lu en sens, soit à la présence d'un V en position 127 de la séquence nucléotidique du gène IFN α -21 et donc que l'allèle T lu en antisens correspond à l'allèle A lu en sens correspondant à un D pour cette position dans la séquence de la protéine correspondante.

En examinant ces résultats par population, on constate que le seul individu hétérozygote AT est issu de la sous-population afro-américaine de la population d'individus.

D) Polymorphisme de type SNP t1155a

L'étape d'élongation ou de miniséquençage est ensuite réalisée comme indiqué dans le tableau suivant :

| Fournisseur | Référence | Réactif | Conc. initiale | Vol. par tube (μ l) | Conc. finale |
|--------------------|---------------------------|--|------------------|--------------------------|--------------------|
| Propre préparation | | Tampon Elongation ¹ (X) | 5 | 1 | 1 |
| Life Technologies | Sur demande | Amorce Miniseq (μ M) G ou H | 10 | 0,5 | 1 |
| AP Biotech | 27-2051 (61,71,81)-01 | ddNTPs ² (μ M) 2 non marqués | 2,5 de chaque | 0,25 | 0,125 de chaque |
| NEN | Nel 472/5 et Nel 492/5 | ddNTPs ² (μ M) 2 marqués Tamra et R110 | 2,5 de chaque | 0,25 | 0,125 de chaque |
| AP Biotech | E79000Z | Thermo-sequenase | 3,2 U/ μ l | 0,125 | 0,4 U/ réaction |
| | | H ₂ O | Qsp 5 μ l | 3,125 | |
| | | PCR digérée | | 10 μ l | |
| | | Vol total | | 15 μ l | |



- 1 Le tampon élongation : Le tampon élongation 5X est composé de Tris-HCl pH 9 à 250 mM, de KCl à 250 mM, de NaCl à 25 mM, de MgCl₂ à 10 mM et de glycérol à 40 %.
- 2 ddNTPs : Pour les ddNTPs, un mélange des 4 bases est réalisé en fonction du polymorphisme étudié. Seulement les 2 bases d'intérêts (T/A) composant le polymorphisme de type SNP
- 5 fonctionnel portent un marquage, soit en Tamra, soit en R110. Le mélange de ddNTPs est composé de :
- 2,5 µM de ddCTP non marqué,
 - 2,5 µM de ddGTP non marqué,
 - 2,5 µM de ddTTP (1,825 µM de ddTTP non marqué et 0,625 µM de ddTTP marqué au Tamra),
 - 10 - 2,5 µM de ddATP (1,825 µM de ddATP non marqué et 0,625 µM de ddATP marqué au R110).

Une fois remplie, la plaque est scellée, centrifugée puis placée dans un thermocycleur pour plaque 384 puits (Tetrad de MJ Research) et subit l'incubation suivante : Cycles d'élongation : 1 min. à 93° C, suivi de 35 cycles

15 composés de 2 étapes (10 sec. à 93° C, 30 sec. à 55° C).

Après la dernière étape dans le thermocycleur, la plaque est directement placée sur un lecteur de fluorescence polarisée de type Analyst® HT de LJJL Biosystems Inc.. La plaque est lue à l'aide du logiciel Criterion Host® en utilisant deux méthodes. La première permet de lire la base

20 marquée en Tamra en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 550-10 nm, émission 580-10 nm) et la seconde permet de lire la base marquée en R110 en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 490-10 nm, émission 520-10 nm). Dans les deux cas, un miroir double dichroïque (R110/Tamra) est

25 utilisé et les autres paramètres de lecture sont :

Z-height : 1,5 mm

Attenuator : out

Temps d'intégration : 100,000 µsec.

Raw data units : counts/sec

30 Switch polarization : by well

Plate settling time : 0 msec

PMT setup : Smart Read (+), sensitivity 2

Dynamic polarizer : emission

Static polarizer : S



Un fichier résultat est alors obtenu contenant les valeurs calculées de mP pour le filtre Tamra et celle pour le filtre R110. Ces valeurs de mP sont calculées à partir des valeurs d'intensité obtenues sur le plan parallèle(//) et sur le plan perpendiculaire (⊥) d'après la formule suivante :

$$5 \quad mP = 1000(// - g\perp)/(// + g\perp).$$

Dans ce calcul, la valeur sur le filtre ⊥ est pondérée d'un facteur g. Celui-ci est un paramètre machine qui doit être déterminé préalablement expérimentalement.

10 Etapes 4) et 5) Interprétation de la lecture et détermination des génotypes.

Les valeurs de mP sont reportées sur un graphe à l'aide du logiciel Excel de Microsoft Inc., et/ou du logiciel Allele Caller® développé par LJI Biosystems Inc..

En abscisse est indiquée la valeur de mP de la base marquée au
15 Tamra, en ordonnée est indiquée la valeur de mP de la base marquée au R110. Une forte valeur de mP indique que la base marquée avec ce fluorophore est incorporée et, inversement, une faible valeur de mP révèle l'absence d'incorporation de cette base.

On obtient jusqu'à 3 groupes homogènes de séquences
20 nucléotidiques ayant des génotypes différents, comme indiqué dans la Figure 13.

L'utilisation du logiciel Allele Caller® permet, une fois le repérage des différents groupes réalisé, d'extraire directement le génotype défini pour chaque individu sous forme d'un tableau.

25 Les séquences des deux amorces de miniséquençage nécessaires pour le génotypage ont été déterminées. Ces amorces sont sélectionnées pour correspondre à une vingtaine voire une trentaine de nucléotides placés juste en amont du site polymorphe. Du fait que le produit de PCR contenant un SNP est un produit d'ADN double brin, le génotypage peut
30 donc se faire soit sur le brin sens soit sur le brin antisens. Les amorces sélectionnées sont fabriquées par Life Technologies Inc.

Les amorces du miniséquençage sont les suivantes :

Amorce sens : (G) : gagaagaaatacagcccttg

Amorce antisens : (H) : gctctgacaacctcccaggc

Le miniséquençage du polymorphisme de type SNP t1155a a d'abord été validé sur 16 échantillons, puis génotypé sur l'ensemble de la population d'individus composée de 239 individus et 7 blancs.

Conditions de miniséquençage testées :

Condition N° 1 :

Amorce sens + ddTTP-R110 + ddATP-Tamra

Condition N° 2 :

10 Amorce sens + ddATP-R110 + ddTTP-Tamra

Condition N° 3 :

Amorce antisens + ddTTP-R110 + ddATP-Tamra

Condition N° 4 :

Amorce antisens + ddATP-R110 + ddTTP-Tamra

15 Ces 4 conditions ont été testées et la condition N° 4 a été retenue pour le génotypage.

B) Résultats

Après la réalisation complète du processus de génotypage, la détermination des génotypes des individus de la population d'individus pour le SNP fonctionnel étudié ici a été réalisée à l'aide du graphe représenté sur la Figure 13.

Ce génotype est en théorie soit homozygote AA, soit hétérozygote AT soit homozygote TT chez les individus testés. En réalité et comme montré ci-dessous, le génotype homozygote TT n'est pas détecté dans la population d'individus.

Les résultats des contrôles, de la répartition des génotypes déterminés dans la population d'individus et le calcul des différentes fréquences alléliques pour ce polymorphisme de type SNP fonctionnel sont présentés dans les tableaux suivants :



64

| Nombre d'individus | | Nombre de blanc | | Pourcentage de réussite |
|--------------------|-----------|-----------------|---------|-------------------------|
| testés | génotypés | Testés | validés | |
| 239 | 236 | 7 | 7 | 98,8 |



| POPULATION | N | % | Fréquence allélique | | | Génotype TT | | Génotype TA | | Génotype AA | |
|-------------------------|-----|------|---------------------|--------|-----|-------------|---|-------------|-----|-------------|------|
| | | | % | 95% IC | | N | % | N | % | N | % |
| Afro Américain | 50 | 20,9 | 1,0 | 0,0 | 3,0 | 0 | 0 | 1 | 2,0 | 48 | 98,0 |
| Amérindien du Sud Ouest | 5 | 2,1 | 0 | - | - | 0 | 0 | 0 | 0 | 5 | 100 |
| Sud Américain (Andes) | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Caribéen | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 9 | 100 |
| Caucasien | 50 | 20,9 | 0 | - | - | 0 | 0 | 0 | 0 | 49 | 100 |
| Chinois | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Grec | 8 | 3,3 | 0 | - | - | 0 | 0 | 0 | 0 | 8 | 100 |
| Ibérien | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Italien | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Japonais | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Mexicain | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Moyen-Orient | 20 | 8,4 | 0 | - | - | 0 | 0 | 0 | 0 | 20 | 100 |
| Individus du Pacifique | 7 | 2,9 | 0 | - | - | 0 | 0 | 0 | 0 | 7 | 100 |
| Indo-Pakistanaï | 9 | 3,8 | 0 | - | - | 0 | 0 | 0 | 0 | 9 | 100 |
| Sud Américain | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Asie du Sud | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Total | 239 | 100 | 0,2 | 0,0 | 0,6 | 0 | 0 | 1 | 0,4 | 235 | 99,6 |



Dans le tableau ci-dessus,

- N représente le nombre d'individus,
- % représente le pourcentage d'individus dans la sous-population spécifique,
- la fréquence allélique représente le pourcentage de l'allèle muté T dans la sous-population spécifique,
- 95 % IC représente l'intervalle minimal et maximal de confiance à 95 %.

Il faut préciser que l'allèle A lu en antisens correspond à l'allèle T lu en sens, soit à la présence d'une cystéine en position 162 de la séquence nucléotidique du gène IFN α -21 et donc que l'allèle T lu en antisens correspond à l'allèle A lu en sens correspondant à un codon stop pour cette position dans la séquence de la protéine correspondante.

En examinant ces résultats par population, on constate que le seul individu hétérozygote AT est issu de la sous-population afro-américaine de la population d'individus.

E) Polymorphisme de type SNP a1204g

L'étape d'élongation ou de miniséquençage est ensuite réalisée comme indiqué dans le tableau suivant :

| Fournisseur | Référence | Réactif | Conc. initiale | Vol. par tube (μ l) | Conc. finale |
|--------------------|---------------------------|--|------------------|--------------------------|--------------------|
| Propre préparation | | Tampon Elongation ¹ (X) | 5 | 1 | 1 |
| Life Technologies | Sur demande | Amorce Miniseq (μ M) I ou J | 10 | 0,5 | 1 |
| AP Biotech | 27-2051 (61,71,81)-01 | ddNTPs ² (μ M) 2 non marqués | 2,5 de chaque | 025 | 0,125 de chaque |
| NEN | Nel 472/5 et Nel 492/5 | ddNTPs ² (μ M) 2 marqués Tamra et R110 | 2,5 de chaque | 0,25 | 0,125 de chaque |
| AP Biotech | E79000Z | Thermo-sequenase | 3,2 U/ μ l | 0,125 | 0,4 U/ réaction |
| | | H2O | Qsp 5 μ l | 3,125 | |
| | | PCR digérée | | 10 μ l | |
| | | Vol total | | 15 μ l | |

- 1 Le tampon élongation : Le tampon élongation 5X est composé de Tris-HCl pH 9 à 250 mM, de KCl à 250 mM, de NaCl à 25 mM, de MgCl₂ à 10 mM et de glycérol à 40 %.
- 2 ddNTPs : Pour les ddNTPs, un mélange des 4 bases est réalisé en fonction du polymorphisme étudié. Seulement les 2 bases d'intérêts (A/G) composant le polymorphisme de type SNP
- 5 fonctionnel portent un marquage, soit en Tamra, soit en R110. Le mélange de ddNTPs est composé de :
- 2,5 µM de ddCTP non marqué,
 - 2,5 µM de ddTTP non marqué,
 - 2,5 µM de ddATP (1,825 µM de ddATP non marqué et 0,625 µM de ddATP marqué au Tamra),

10 - 2,5 µM de ddGTP (1,825 µM de ddGTP non marqué et 0,625 µM de ddGTP marqué au R110).

Une fois remplie, la plaque est scellée, centrifugée puis placée dans un thermocycleur pour plaque 384 puits (Tetrad de MJ Research) et subit l'incubation suivante : Cycles d'élongation : 1 min. à 93° C, suivi de 35 cycles

15 composés de 2 étapes (10 sec. à 93° C, 30 sec. à 55° C).

Après la dernière étape dans le thermocycleur, la plaque est directement placée sur un lecteur de fluorescence polarisée de type Analyst® HT de LJJL Biosystems Inc.. La plaque est lue à l'aide du logiciel Criterion Host® en utilisant deux méthodes. La première permet de lire la base

20 marquée en Tamra en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 550-10 nm, émission 580-10 nm) et la seconde permet de lire la base marquée en R110 en utilisant les filtres d'excitation et d'émission spécifiques de ce fluorophore (excitation 490-10 nm, émission 520-10 nm). Dans les deux cas, un miroir double dichroïque (R110/Tamra) est

25 utilisé et les autres paramètres de lecture sont :

Z-height : 1,5 mm

Attenuator : out

Temps d'intégration : 100,000 µsec.

Raw data units : counts/sec

30 Switch polarization : by well

Plate settling time : 0 msec

PMT setup : Smart Read (+), sensitivity 2

Dynamic polarizer : emission

Static polarizer : S



Un fichier résultat est alors obtenu contenant les valeurs calculées de mP pour le filtre Tamra et celle pour le filtre R110. Ces valeurs de mP sont calculées à partir des valeurs d'intensité obtenues sur le plan parallèle(//) et sur le plan perpendiculaire (⊥) d'après la formule suivante :

5
$$mP = 1000(// - g\perp)/(// + g\perp).$$

Dans ce calcul, la valeur sur le filtre ⊥ est pondérée d'un facteur g. Celui-ci est un paramètre machine qui doit être déterminé préalablement expérimentalement.

10 Etapes 4) et 5) Interprétation de la lecture et détermination des génotypes.

Les valeurs de mP sont reportées sur un graphe à l'aide du logiciel Excel de Microsoft Inc., et/ou du logiciel Allele Caller® développé par LJI Biosystems Inc..

15 En abscisse est indiquée la valeur de mP de la base marquée au Tamra, en ordonnée est indiquée la valeur de mP de la base marquée au R110. Une forte valeur de mP indique que la base marquée avec ce fluorophore est incorporée et, inversement, une faible valeur de mP révèle l'absence d'incorporation de cette base.

20 On obtient jusqu'à 3 groupes homogènes de séquences nucléotidiques ayant des génotypes différents, comme indiqué dans la Figure 14.

L'utilisation du logiciel Allele Caller® permet, une fois le repérage des différents groupes réalisé, d'extraire directement le génotype défini pour chaque individu sous forme d'un tableau.

25 Les séquences des deux amorces de miniséquençage nécessaires pour le génotypage ont été déterminées. Ces amorces sont sélectionnées pour correspondre à une vingtaine voire une trentaine de nucléotides placés juste en amont du site polymorphe. Du fait que le produit de PCR contenant un SNP est un produit d'ADN double brin, le génotypage peut
30 donc se faire soit sur le brin sens soit sur le brin antisens. Les amorces sélectionnées sont fabriquées par Life Technologies Inc.

Les amorces du miniséquençage sont les suivantes :

Amorce sens : (I) : tgagatccttctctttatca

Amorce antisens : (J) : taatcttcttgaaaaattt

Le miniséquençage du polymorphisme de type SNP a1204g a d'abord été validé sur 16 échantillons, puis génotypé sur l'ensemble de la population d'individus composée de 239 individus et 7 blancs.

Conditions de miniséquençage testées :

Condition N° 1 :

Amorce sens + ddATP-R110 + ddGTP-Tamra

Condition N° 2 :

10 Amorce sens + ddGTP-R110 + ddATP-Tamra

Condition N° 3 :

Amorce antisens + ddTTP-R110 + ddCTP-Tamra

Condition N° 4 :

Amorce antisens + ddCTP-R110 + ddTTP-Tamra

15 Ces 4 conditions ont été testées et la condition N° 2 a été retenue pour le génotypage.

B) Résultats

Après la réalisation complète du processus de génotypage, la détermination des génotypes des individus de la population d'individus pour le SNP fonctionnel étudié ici a été réalisée à l'aide du graphe représenté sur la Figure 14.

Ce génotype est en théorie soit homozygote sauvage AA, soit hétérozygote AG soit homozygote muté GG chez les individus testés. En réalité et comme montré ci-dessous, le génotype homozygote GG n'est pas détecté dans la population d'individus.

Les résultats des contrôles, de la répartition des génotypes déterminés dans la population d'individus et le calcul des différentes fréquences alléliques pour ce polymorphisme de type SNP fonctionnel sont présentés dans les tableaux suivants :



70

| Nombre d'individus | | Nombre de blanc | | Pourcentage de réussite |
|--------------------|-----------|-----------------|---------|-------------------------|
| testés | génotypés | testés | validés | |
| 239 | 238 | 7 | 7 | 94,7 |

| POPULATION | N | % | Fréquence allélique | | | Génotype GG | | Génotype GA | | Génotype AA | |
|-------------------------|-----|------|---------------------|--------|------|-------------|-----|-------------|------|-------------|------|
| | | | % | 95% IC | | N | % | N | % | N | % |
| Afro Américain | 50 | 20,9 | 0 | - | - | 0 | 0 | 0 | 0,0 | 50 | 100 |
| Amérindien du Sud Ouest | 5 | 2,1 | 10,0 | 0,0 | 28,6 | 0 | 0 | 1 | 20,0 | 4 | 80,0 |
| Sud Américain (Andes) | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 9 | 100 |
| Caribéen | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Caucasien | 50 | 20,9 | 0 | - | - | 0 | 0 | 0 | 0 | 50 | 100 |
| Chinois | 10 | 4,2 | 15,0 | 0,0 | 30,6 | 0 | 0 | 3 | 30,0 | 7 | 70 |
| Grec | 8 | 3,3 | 0 | - | - | 0 | 0 | 0 | 0 | 8 | 100 |
| Ibérien | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Italien | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Japonais | 10 | 4,2 | 20,0 | 2,5 | 37,5 | 0 | 0 | 4 | 40,0 | 6 | 60,0 |
| Mexicain | 10 | 4,2 | 10,0 | 0,0 | 23,1 | 0 | 0 | 2 | 20,0 | 8 | 80,0 |
| Moyen-Orient | 20 | 8,4 | 0 | - | - | 0 | 0 | 0 | 0 | 20 | 100 |
| Individus du Pacifique | 7 | 2,9 | 0 | - | - | 0 | 0 | 0 | 0 | 7 | 100 |
| Indo-Pakistanaïs | 9 | 3,8 | 11,1 | 0,0 | 25,6 | 0 | 0 | 2 | 22,2 | 7 | 77,8 |
| Sud Américain | 10 | 4,2 | 5,0 | 0,0 | 14,6 | 0 | 0 | 1 | 10 | 9 | 90 |
| Asie du Sud | 10 | 4,2 | 0 | - | - | 0 | 0 | 0 | 0 | 10 | 100 |
| Total | 239 | 100 | 2,7 | 1,3 | 4,2 | 0 | 0,0 | 13 | 5,5 | 225 | 94,5 |



Dans le tableau ci-dessus,

- N représente le nombre d'individus,
 - % représente le pourcentage d'individus dans la sous-population spécifique,
 - la fréquence allélique représente le pourcentage de l'allèle muté dans la sous-population spécifique,*
- 5
- 95 % IC représente l'intervalle minimal et maximal de confiance à 95 %.

En examinant ces résultats par population, on constate que les individus hétérozygotes mutés AG sont issus de 6 sous-populations : amérindienne du sud-ouest, chinoise, japonaise, mexicaine, indo-pakistanaise

10 et sud-américaine de la population d'individus.

REVENDEICATIONS**1. Polynucléotide isolé comprenant :**

- 5 a) une séquence nucléotidique ayant au moins 80 % d'identité avec la séquence ID SEQ N° 1 ou sa séquence codante, sous réserve que cette séquence nucléotidique comporte au moins l'un des polymorphismes de type SNP suivants :
- c973a,
 - g1011c,
 - 10 - t1049a,
 - t1155a,
 - a1204g; ou
- b) une séquence nucléotidique complémentaire à une séquence nucléotidique sous a).

15 2. Polynucléotide isolé comprenant :

- a) la séquence nucléotidique ID SEQ N° 1 ou sa séquence codante sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP suivants :
- c973a,
 - 20 - g1011c,
 - t1049a,
 - t1155a,
 - a1204g; ou
- b) une séquence nucléotidique complémentaire à une séquence nucléotidique
- 25 sous a).

3. Polynucléotide selon la revendication 2, caractérisé en ce qu'il consiste en la séquence ID SEQ N° 1 ou sa séquence codante, sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP suivants :

- 30 - c973a,
- g1011c,
 - t1049a,



REVENDEICATIONS

1. Polynucléotide isolé comprenant :

- a) une séquence nucléotidique ayant au moins 80 % d'identité avec la séquence ID SEQ N° 1 ou sa séquence codante, sous réserve que cette séquence nucléotidique comporte au moins l'un des polymorphismes de type SNP suivants :
 - c973a,
 - g1011c,
 - t1049a,
 - t1155a,
 - a1204g; ou
- b) une séquence nucléotidique complémentaire à une séquence nucléotidique sous a).

2. Polynucléotide isolé comprenant :

- a) la séquence nucléotidique ID SEQ N° 1 ou sa séquence codante sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP suivants :
 - c973a,
 - g1011c,
 - t1049a,
 - t1155a,
 - a1204g; ou
- b) une séquence nucléotidique complémentaire à une séquence nucléotidique sous a).

3. Polynucléotide selon la revendication 2, caractérisé en ce qu'il consiste en la séquence ID SEQ N° 1 ou sa séquence codante, sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP suivants :

- c973a,
- g1011c,
- t1049a,



- t1155a,
- a1204g.

4. Polynucléotide selon l'une quelconque des revendications 1 à 3, caractérisé en ce que la séquence nucléotidique a) comporte un seul polymorphisme de type SNP choisi dans le groupe constitué par :

- c973a,
 - g1011c,
 - t1049a,
 - t1155a, et
- 10 - a1204g.

5. Polynucléotide selon l'une quelconque des revendications 1 à 4, caractérisé en ce qu'il code pour un polypeptide comprenant :

- a) la séquence d'acides aminés ID SEQ N° 2, ou
 - b) la séquence d'acides aminés comprenant les acides aminés compris entre 24
- 15 et 189 de la séquence d'acides aminés ID SEQ N° 2,
- sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

- Q102K,
 - Q114H,
- 20 - V127D,
- C162stop,
 - K179E.

6. Polynucléotide selon la revendication 5, caractérisé en ce qu'il code pour un polypeptide comportant un seul polymorphisme de type SNP

25 choisi dans le groupe constitué par :

- Q102K,
 - Q114H,
 - V127D,
 - C162stop,
- 30 - K179E.

7. Polynucléotide selon l'une quelconque des revendications 1 à 6, caractérisé en ce qu'il comprend une molécule d'ADN ou d'ARN.

- t1155a,
- a1204g.

4. Polynucléotide selon l'une quelconque des revendications 1 à 3, caractérisé en ce que la séquence nucléotidique a) comporte un seul polymorphisme de type SNP choisi dans le groupe constitué par :

- c973a,
- g1011c,
- t1049a,
- t1155a, et
- a1204g.

5. Polynucléotide selon l'une quelconque des revendications 1 à 4, caractérisé en ce qu'il code pour un polypeptide comprenant :

- a) la séquence d'acides aminés ID SEQ N° 2, ou
 - b) la séquence d'acides aminés comprenant les acides aminés compris entre 24 et 189 de la séquence d'acides aminés ID SEQ N° 2,
- sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

- Q102K,
- Q114H,
- V127D,
- C162stop,
- K179E.

6. Polynucléotide selon la revendication 5, caractérisé en ce qu'il code pour un polypeptide comportant un seul polymorphisme de type SNP choisi dans le groupe constitué par :

- Q102K,
- Q114H,
- V127D,
- C162stop,
- K179E.

7. Polynucléotide selon l'une quelconque des revendications 1 à 6, caractérisé en ce qu'il comprend une molécule d'ADN ou d'ARN.

8. Utilisation de tout ou partie d'un polynucléotide selon l'une quelconque des revendications 1 à 7, pour identifier, hybrider et/ou amplifier tout ou partie d'un polynucléotide consistant en la séquence ID SEQ N° 1 ou sa séquence codante, sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP suivants:
- c973a,
 - g1011c,
 - t1049a,
 - t1155a,
 - a1204g.
9. Utilisation de tout ou partie d'un polynucléotide selon l'une quelconque des revendications 1 à 7, comme outil de génotypage.
10. Procédé de détermination de la fréquence du polymorphisme de type SNP d'un polynucléotide selon l'une quelconque des revendications 1 à 7, dans lequel on procède à un génotypage chez un individu ou dans une population d'individus.
11. Procédé de détermination selon la revendication 10, dans lequel le génotypage est effectué par miniséquençage.
12. Procédé de détermination selon la revendication 11, dans lequel le miniséquençage est réalisé avec les amorces sens et antisens correspondant respectivement aux séquences nucléotidique ID SEQ N° 3 et ID SEQ N° 4.
13. Utilisation d'un polynucléotide selon l'une quelconque des revendications 1 à 6, pour la recherche d'une variation de séquence dans la séquence nucléotidique du gène IFN α -21 chez un individu.
14. Vecteur recombinant comprenant un polynucléotide selon l'une quelconque des revendications 1 à 7.
15. Cellule hôte comprenant un vecteur recombinant selon la revendication 14.
16. Procédé de préparation d'un polypeptide, caractérisé en ce qu'une cellule hôte selon la revendication 15 est cultivée et ledit polypeptide est isolé du milieu de culture.

8. Utilisation de tout ou partie d'un polynucléotide selon l'une quelconque des revendications 1 à 7, pour identifier, hybrider et/ou amplifier tout ou partie d'un polynucléotide consistant en la séquence ID SEQ N° 1 ou sa séquence codante, sous réserve que chacune de ces séquences comporte au moins l'un des polymorphismes de type SNP suivants:

- c973a,
- g1011c,
- t1049a,
- t1155a,
- a1204g.

9. Utilisation de tout ou partie d'un polynucléotide selon l'une quelconque des revendications 1 à 7, comme outil de génotypage.

10. Procédé de détermination de la fréquence du polymorphisme de type SNP d'un polynucléotide selon l'une quelconque des revendications 1 à 7, dans lequel on procède à un génotypage chez un individu ou dans une population d'individus.

11. Procédé de détermination selon la revendication 10, dans lequel le génotypage est effectué par miniséquençage.

12. Procédé de détermination selon la revendication 11, dans lequel le miniséquençage est réalisé avec les amorces sens et antisens correspondant respectivement aux séquences nucléotidique ID SEQ N° 3 et ID SEQ N° 4.

13. Utilisation d'un polynucléotide selon l'une quelconque des revendications 1 à 7, pour la recherche d'une variation de séquence dans la séquence nucléotidique du gène IFN α -21 chez un individu.

14. Vecteur recombinant comprenant un polynucléotide selon l'une quelconque des revendications 1 à 7.

15. Cellule hôte comprenant un vecteur recombinant selon la revendication 14.

16. Procédé de préparation d'un polypeptide, caractérisé en ce qu'une cellule hôte selon la revendication 15 est cultivée et ledit polypeptide est isolé du milieu de culture.



17. Polypeptide isolé comprenant une séquence d'acides aminés ayant au moins 80 % d'identité avec :

a) la séquence d'acides aminés ID SEQ N° 2, ou avec

b) la séquence d'acides aminés comportant les acides aminés compris entre 24

5 et 189 de la séquence d'acides aminés ID SEQ N° 2,

sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

- Q102K,

- Q114H,

10 - V127D,

- C162stop,

- K179E.

18. Polypeptide selon la revendication 17, caractérisé en ce qu'il comprend :

15 a) la séquence d'acides aminés ID SEQ N° 2, ou

b) la séquence d'acides aminés comportant les acides aminés compris entre 24 et 189 de la séquence d'acides aminés ID SEQ N° 2,

sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

20 - Q102K,

- Q114H,

- V127D,

- C162stop,

- K179E.

25 19. Polypeptide selon l'une quelconque des revendications 17 à 18, caractérisé en ce qu'il consiste en :

a) la séquence d'acides aminés ID SEQ N° 2, ou

b) la séquence d'acides aminés comportant les acides aminés compris entre 24 et 189 de la séquence d'acides aminés ID SEQ N° 2,

30 sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

- Q102K,



17. Polypeptide isolé comprenant une séquence d'acides aminés ayant au moins 80 % d'identité avec :

- a) la séquence d'acides aminés ID SEQ N° 2, ou avec
- b) la séquence d'acides aminés comportant les acides aminés compris entre 24 et 189 de la séquence d'acides aminés ID SEQ N° 2,

sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

- Q102K,
- Q114H,
- V127D,
- C162stop,
- K179E.

18. Polypeptide selon la revendication 17, caractérisé en ce qu'il comprend :

- a) la séquence d'acides aminés ID SEQ N° 2, ou
- b) la séquence d'acides aminés comportant les acides aminés compris entre 24 et 189 de la séquence d'acides aminés ID SEQ N° 2,

sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

- Q102K,
- Q114H,
- V127D,
- C162stop,
- K179E.

19. Polypeptide selon l'une quelconque des revendications 17 à 18, caractérisé en ce qu'il consiste en :

- a) la séquence d'acides aminés ID SEQ N° 2, ou
- b) la séquence d'acides aminés comportant les acides aminés compris entre 24 et 189 de la séquence d'acides aminés ID SEQ N° 2,

sous réserve que chacune des séquences d'acides aminés sous a) et b) comporte au moins l'un des polymorphismes de type SNP suivants :

- Q102K,



- Q114H,
- V127D,
- C162stop,
- K179E.

5 20. Polypeptide selon l'une quelconque des revendications 17 à 19, caractérisé en ce que chacune des séquences d'acides aminés sous a) et b) comporte un seul polymorphisme de type SNP choisi dans le groupe constitué par:

- Q102K,
- 10 - Q114H,
- V127D,
- C162stop, et
- K179E.

21. Procédé d'obtention d'un anticorps immunospécifique,
15 caractérisé en ce qu'il est obtenu par immunisation d'un animal avec un polypeptide selon l'une quelconque des revendications 17 à 20.

21. Anticorps immunospécifique pour un polypeptide selon l'une quelconque des revendications 17 à 20.

22. Procédé d'identification d'un agent activateur ou inhibiteur d'un
20 polypeptide selon l'une quelconque des revendications 17 à 20, comprenant :
a) la mise en présence de cellules hôtes selon la revendication 15 avec un agent à tester, et
b) la détermination de l'effet activateur ou inhibiteur généré par l'agent à tester.

23. Agent activateur ou inhibiteur, caractérisé en ce qu'il est
25 susceptible d'être identifié par la méthode selon la revendication 22.

24. Méthode pour l'identification d'un agent activé ou inhibé par un polypeptide selon l'une quelconque des revendications 17 à 20, comprenant :
a) la mise en présence de cellules hôtes selon la revendication 15 avec un agent à tester, et
30 b) la détermination de l'effet activateur ou inhibiteur généré par un polypeptide sur l'agent à tester.

25. Agent activé ou inhibé, caractérisé en ce qu'il est susceptible



- Q114H,
- V127D,
- C162stop,
- K179E.

20. Polypeptide selon l'une quelconque des revendications 17 à 19, caractérisé en ce que chacune des séquences d'acides aminés sous a) et b) comporte un seul polymorphisme de type SNP choisi dans le groupe constitué par:

- Q102K,
- Q114H,
- V127D,
- C162stop, et
- K179E.

21. Procédé d'obtention d'un anticorps immunosécifique, caractérisé en ce qu'il est obtenu par immunisation d'un animal avec un polypeptide selon l'une quelconque des revendications 17 à 20.

22. Anticorps immunosécifique pour un polypeptide selon l'une quelconque des revendications 17 à 20.

23. Procédé d'identification d'un agent activateur ou inhibiteur d'un polypeptide selon l'une quelconque des revendications 17 à 20, comprenant :

- a) la mise en présence de cellules hôtes selon la revendication 15 avec un agent à tester, et
- b) la détermination de l'effet activateur ou inhibiteur généré par l'agent à tester.

24. Méthode pour l'identification d'un agent activé ou inhibé par un polypeptide selon l'une quelconque des revendications 17 à 20, comprenant :

- a) la mise en présence de cellules hôtes selon la revendication 15 avec un agent à tester, et
- b) la détermination de l'effet activateur ou inhibiteur généré par un polypeptide sur l'agent à tester.

25. Procédé de détection de l'expression et/ou de l'activité d'un polypeptide selon l'une quelconque des revendications 17 à 20, comprenant :

- a) la détection de la présence ou de l'absence d'un polynucléotide selon l'une

d'être identifié par la méthode selon la revendication 24.

26. Procédé de détection de l'expression et/ou de l'activité d'un polypeptide selon l'une quelconque des revendications 17 à 20, comprenant :

- 5 a) la détection de la présence ou de l'absence d'un polynucléotide selon l'une quelconque des revendications 1 à 7 dans le génome du sujet, et/ou
- b) la détection de la présence, de l'absence et/ou d'une concentration prédéterminée d'un polypeptide selon l'une quelconque des revendications 17 à 20, dans un échantillon biologique du sujet.

27. Médicament comprenant à titre de principe actif au moins un
10 polypeptide selon l'une quelconque des revendications 17 à 20.

28. Utilisation d'un polypeptide selon l'une quelconque des revendications 17 à 20, pour la préparation d'un médicament destiné à la prévention ou le traitement de l'une des maladies choisies dans le groupe constitué par les cancers comme, les carcinomes, les mélanomes, les
15 lymphomes, les myélomes, les tumeurs, les leucémies et les cancers du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que celles qui ne sont pas liées au système immunitaire comme, l'obésité, les maladies infectieuses en particulier les infections virales comme les hépatites B et C et le SIDA, les pneumonies, les colites ulcéraives,
20 les maladies du système nerveux central comme, la schizophrénie et la dépression, le rejet de greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie, les allergies, l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les maladies et les désordres auto-immuns, les verrues génitales ou vénériennes,
25 les désordres gastro-intestinaux et les désordres liés aux traitements par chimiothérapie.

29. Utilisation d'un polypeptide selon la revendication 28, pour la préparation d'un médicament destiné à la prévention ou le traitement de l'une des maladies choisies dans le groupe constitué par les hépatites chroniques B
30 et C, les leucémies telles que la leucémie à trichloroleucocytes et la leucémie myéloïde chronique, les myélomes multiples, des lymphomes folliculaires, les tumeurs carcinoïdes, les mélanomes malins, les carcinomes rénaux

- quelconque des revendications 1 à 7 dans le génome du sujet, et/ou
- b) la détection de la présence, de l'absence et/ou d'une concentration prédéterminée d'un polypeptide selon l'une quelconque des revendications 17 à 20, dans un échantillon biologique du sujet.

26. Médicament comprenant à titre de principe actif au moins un polypeptide selon l'une quelconque des revendications 17 à 20.

27. Utilisation d'un polypeptide selon l'une quelconque des revendications 17 à 20, pour la préparation d'un médicament destiné à la prévention ou le traitement de l'une des maladies choisies dans le groupe constitué par les cancers comme, les carcinomes, les mélanomes, les lymphomes, les myélomes, les tumeurs, les leucémies et les cancers du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que celles qui ne sont pas liées au système immunitaire comme, l'obésité, les maladies infectieuses en particulier les infections virales comme les hépatites B et C et le SIDA, les pneumonies, les colites ulcéraives, les maladies du système nerveux central comme, la schizophrénie et la dépression, le rejet de greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie, les allergies, l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les maladies et les désordres auto-immuns, les verrues génitales ou vénériennes, les désordres gastro-intestinaux et les désordres liés aux traitements par chimiothérapie.

28. Utilisation d'un polypeptide selon l'une quelconque des revendications 17 à 20, pour la préparation d'un médicament destiné à la prévention ou le traitement de l'une des maladies choisies dans le groupe constitué par les hépatites chroniques B et C, les leucémies telles que la leucémie à trichoroleucocytes et la leucémie myéloïde chronique, les myélomes multiples, des lymphomes folliculaires, les tumeurs carcinoïdes, les mélanomes malins, les carcinomes rénaux métastasés, la maladie d'Alzheimer, la maladie de Parkinson ainsi que les tumeurs qui apparaissent à la suite d'un déficit immunitaire, tel que le sarcome de Kaposi dans le cas du SIDA.

29. Médicament comprenant à titre de principe actif au moins un

métastasés, la maladie d'Alzheimer, la maladie de Parkinson ainsi que les tumeurs qui apparaissent à la suite d'un déficit immunitaire, tel que le sarcome de Kaposi dans le cas du SIDA.

30. Médicament comprenant à titre de principe actif au moins un
5 polynucléotide selon l'une quelconque des revendications 17 à 20, un vecteur recombinant selon la revendication 14, une cellule hôte selon la revendication 15, un anticorps selon la revendication 21 et/ou un agent activateur et/ou inhibiteur selon la revendication 23.

31. Utilisation d'un polynucléotide selon l'une quelconque des
10 revendications 1 à 7, d'un vecteur recombinant selon la revendication 14, d'une cellule hôte selon la revendication 15, d'un anticorps selon la revendication 21 et/ou d'un agent activateur et/ou inhibiteur selon la revendication 23, pour la préparation d'un médicament destiné à la prévention ou le traitement de l'une des maladies choisies dans le groupe constitué par les cancers comme, par
15 exemple, les carcinomes, les mélanomes, les lymphomes, les myélomes, les tumeurs, les leucémies et les cancers du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que celles qui ne sont pas liées au système immunitaire comme, par exemple, l'obésité, les maladies infectieuses en particulier les infections virales comme les hépatites B
20 et C et le SIDA, les pneumonies, les colites ulcéraives, les maladies du système nerveux central comme, par exemple, la schizophrénie et la dépression, le rejet de greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie, les allergies, l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les
25 maladies et les désordres auto-immuns, les verrues génitales ou vénériennes, les désordres gastro-intestinaux et les désordres liés aux traitements par chimiothérapie.

32. Utilisation d'un polynucléotide, d'un vecteur recombinant, d'une cellule hôte, d'un anticorps et/ou d'un agent activateur et/ou inhibiteur
30 selon la revendication 31, pour la préparation d'un médicament destiné à la prévention ou le traitement de l'une des maladies choisies dans le groupe constitué par les hépatites chroniques B et C, les leucémies telles que la

polynucléotide selon l'une quelconque des revendications 1 à 7, un vecteur recombinant selon la revendication 14, une cellule hôte selon la revendication 15 et/ou un anticorps selon la revendication 22.

30. Utilisation d'un polynucléotide selon l'une quelconque des revendications 1 à 7, d'un vecteur recombinant selon la revendication 14, d'une cellule hôte selon la revendication 15 et/ou d'un anticorps selon la revendication 22, pour la préparation d'un médicament destiné à la prévention ou le traitement de l'une des maladies choisies dans le groupe constitué par les cancers comme, par exemple, les carcinomes, les mélanomes, les lymphomes, les myélomes, les tumeurs, les leucémies et les cancers du foie, du cou, de la tête et des reins, les maladies cardiovasculaires, les maladies métaboliques telles que celles qui ne sont pas liées au système immunitaire comme, par exemple, l'obésité, les maladies infectieuses en particulier les infections virales comme les hépatites B et C et le SIDA, les pneumonies, les colites ulcéraives, les maladies du système nerveux central comme, par exemple, la schizophrénie et la dépression, le rejet de greffe de tissus ou d'organes, la cicatrisation de blessures, l'anémie, les allergies, l'asthme, les scléroses multiples, l'ostéoporose, le psoriasis, l'arthrite rhumatoïde, la maladie de Crohn, les maladies et les désordres auto-immuns, les verrues génitales ou vénériennes, les désordres gastro-intestinaux et les désordres liés aux traitements par chimiothérapie.

31. Utilisation d'un polynucléotide selon l'une quelconque des revendications 1 à 7, d'un vecteur recombinant selon la revendication 14, d'une cellule hôte selon la revendication 15 et/ou d'un anticorps selon la revendication 22, pour la préparation d'un médicament destiné à la prévention ou le traitement de l'une des maladies choisies dans le groupe constitué par les hépatites chroniques B et C, les leucémies telles que la leucémie à trichloroleucocytes et la leucémie myéloïde chronique, les myélomes multiples, des lymphomes folliculaires, les tumeurs carcinoïdes, les mélanomes malins, les carcinomes rénaux métastasés, la maladie d'Alzheimer, la maladie de Parkinson ainsi que les tumeurs qui apparaissent à la suite d'un déficit immunitaire, tel que le sarcome de Kaposi dans le cas du SIDA.

leucémie à trichloroleucocytes et la leucémie myéloïde chronique, les myélomes multiples, des lymphomes folliculaires, les tumeurs carcinoïdes, les mélanomes malins, les carcinomes rénaux métastasés, la maladie d'Alzheimer, la maladie de Parkinson ainsi que les tumeurs qui apparaissent à la suite d'un

5 déficit immunitaire, tel que le sarcome de Kaposi dans le cas du SIDA.

33. Composition pharmaceutique renfermant à titre de principe actif au moins un polypeptide selon l'une quelconque des revendications 17 à 20, un polynucléotide selon l'une quelconque des revendications 1 à 7, un vecteur recombinant selon la revendication 14, une cellule hôte selon la

10 revendication 15, un anticorps selon la revendication 21 et/ou un agent activateur et/ou inhibiteur selon la revendication 23, ainsi qu'un excipient pharmaceutiquement acceptable.

34. Composition de diagnostic renfermant à titre de principe actif au moins un polypeptide selon l'une quelconque des revendications 17 à 20, un

15 polynucléotide selon l'une quelconque des revendications 1 à 7, un vecteur recombinant selon la revendication 14, une cellule hôte selon la revendication 15, un anticorps selon la revendication 21 et/ou un agent activateur et/ou inhibiteur selon la revendication 23, ainsi qu'un excipient approprié pharmaceutiquement acceptable.

20 35. Utilisation :

- a) d'une quantité thérapeutiquement efficace d'un agent activateur selon la revendication 23, et/ou
- b) d'une quantité thérapeutiquement efficace d'un polypeptide selon l'une quelconque des revendications 17 à 20, et/ou
- 25 c) d'un polynucléotide selon l'une quelconque des revendications 1 à 7, et/ou
- d) d'une cellule hôte selon la revendication 15, cette cellule provenant du sujet à traiter,

pour préparer un médicament destiné à augmenter l'expression ou l'activité chez un sujet, d'un polypeptide selon l'une quelconque des revendications 17 à

30 20.

36. Utilisation

- a) d'une quantité thérapeutiquement efficace d'un agent inhibiteur selon la

32. Composition pharmaceutique renfermant à titre de principe actif au moins un polypeptide selon l'une quelconque des revendications 17 à 20, un polynucléotide selon l'une quelconque des revendications 1 à 7, un vecteur recombinant selon la revendication 14, une cellule hôte selon la revendication 15 et/ou un anticorps selon la revendication 22, ainsi qu'un excipient pharmaceutiquement acceptable.

33. Composition de diagnostic renfermant à titre de principe actif au moins un polypeptide selon l'une quelconque des revendications 17 à 20, un polynucléotide selon l'une quelconque des revendications 1 à 7, un vecteur recombinant selon la revendication 14, une cellule hôte selon la revendication 15 et/ou un anticorps selon la revendication 22, ainsi qu'un excipient approprié pharmaceutiquement acceptable.

34. Utilisation :

- a) d'une quantité thérapeutiquement efficace d'un polypeptide selon l'une quelconque des revendications 17 à 20, et/ou
- b) d'un polynucléotide selon l'une quelconque des revendications 1 à 7, et/ou
- c) d'une cellule hôte selon la revendication 15, cette cellule provenant du sujet à traiter,

pour préparer un médicament destiné à augmenter l'expression ou l'activité chez un sujet, d'un polypeptide selon l'une quelconque des revendications 17 à 20.

35. Utilisation :

- a) d'une quantité thérapeutiquement efficace d'un anticorps selon la revendication 22, et/ou
- b) d'un polynucléotide permettant d'inhiber l'expression d'un polynucléotide selon l'une quelconque des revendications 1 à 7,

pour préparer un médicament destiné à diminuer l'expression ou l'activité chez un sujet d'un polypeptide selon l'une quelconque des revendications 17 à 20.



- revendication 23, et/ou
- b) d'une quantité thérapeutiquement efficace d'un anticorps selon la revendication 21, et/ou
 - c) d'un polynucléotide permettant d'inhiber l'expression d'un polynucléotide selon
- 5 l'une quelconque des revendications 1 à 7,
- pour préparer un médicament destiné à diminuer l'expression ou l'activité chez un sujet d'un polypeptide selon l'une quelconque des revendications 17 à 20.

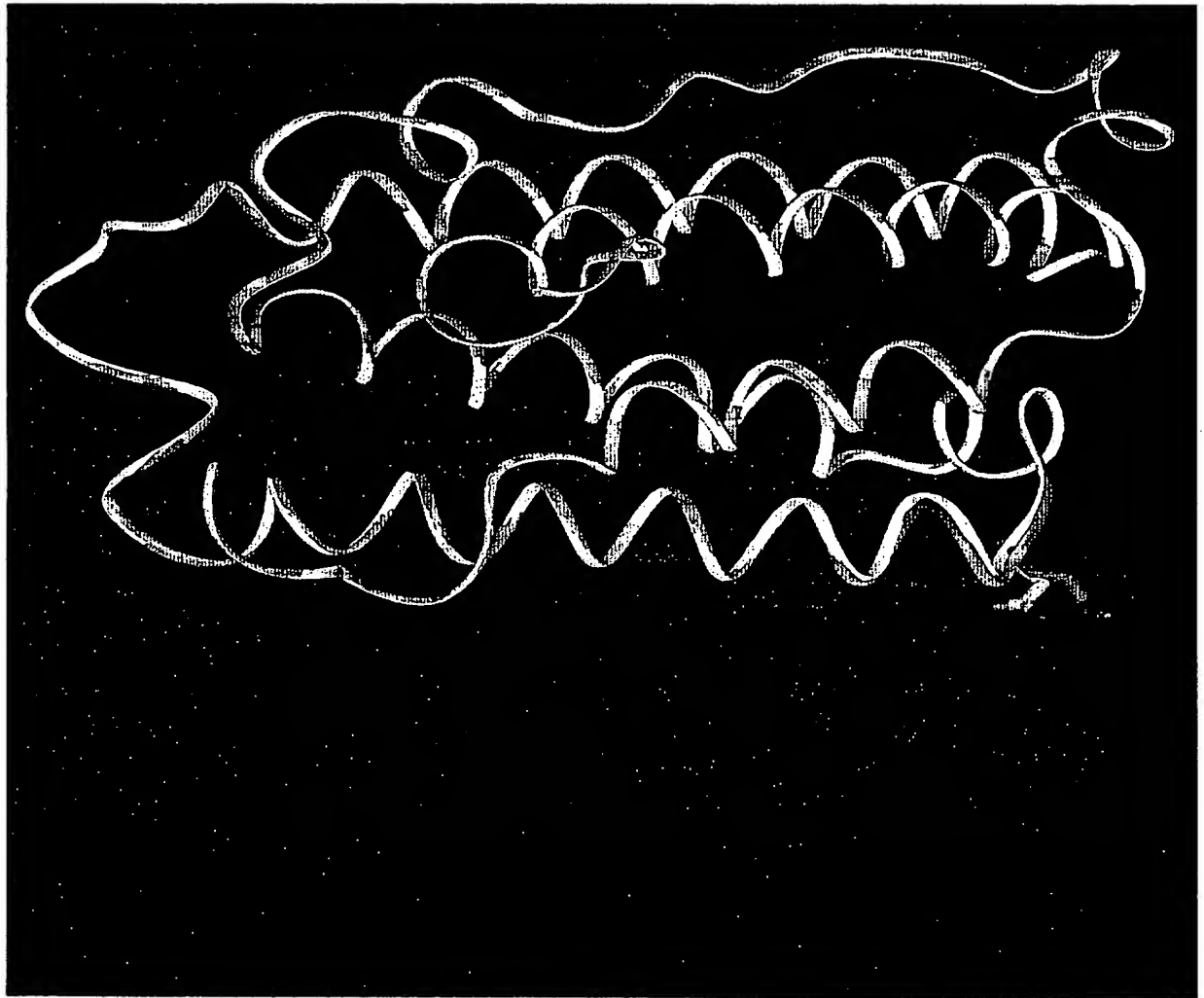


Figure 1

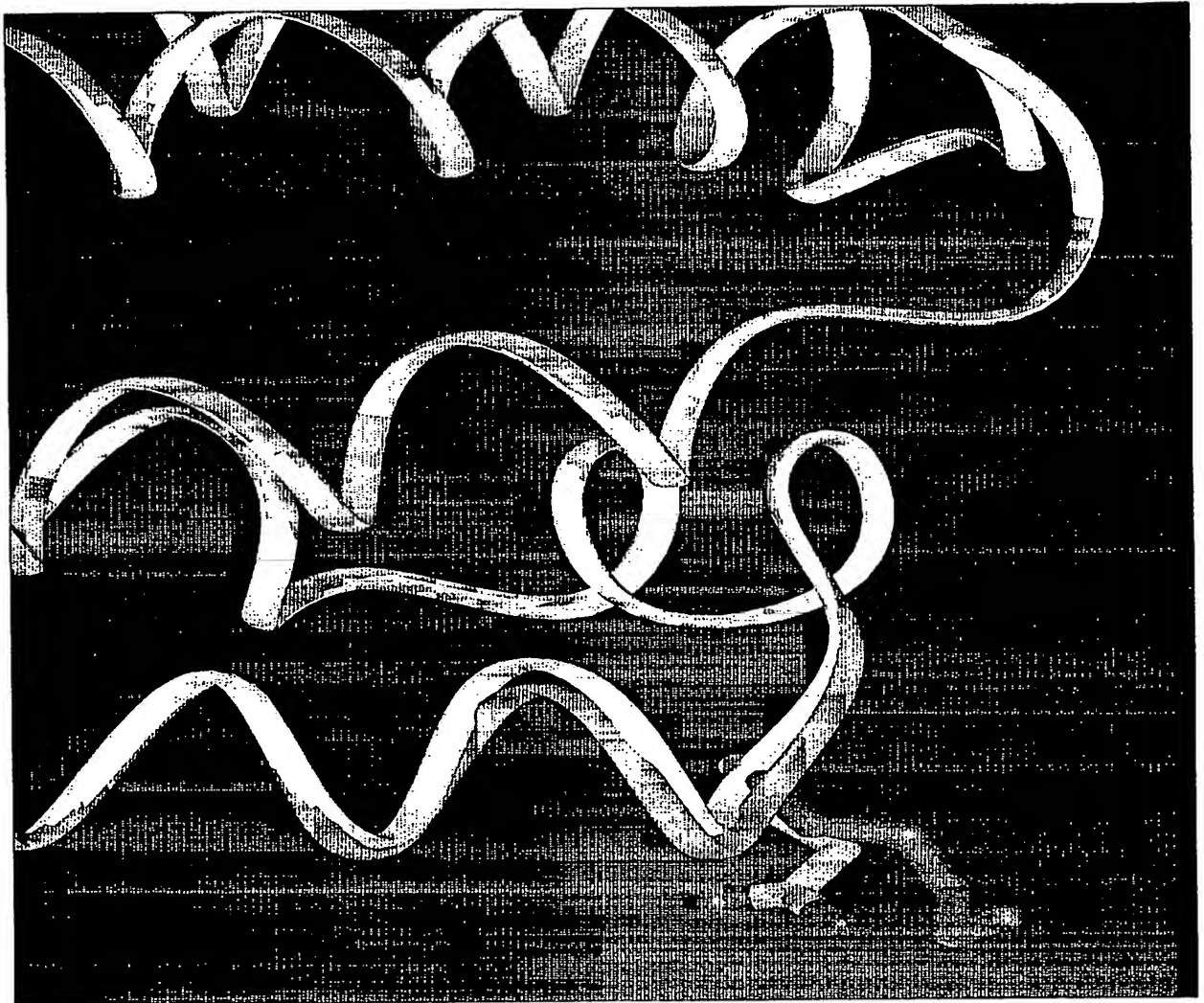


Figure 2

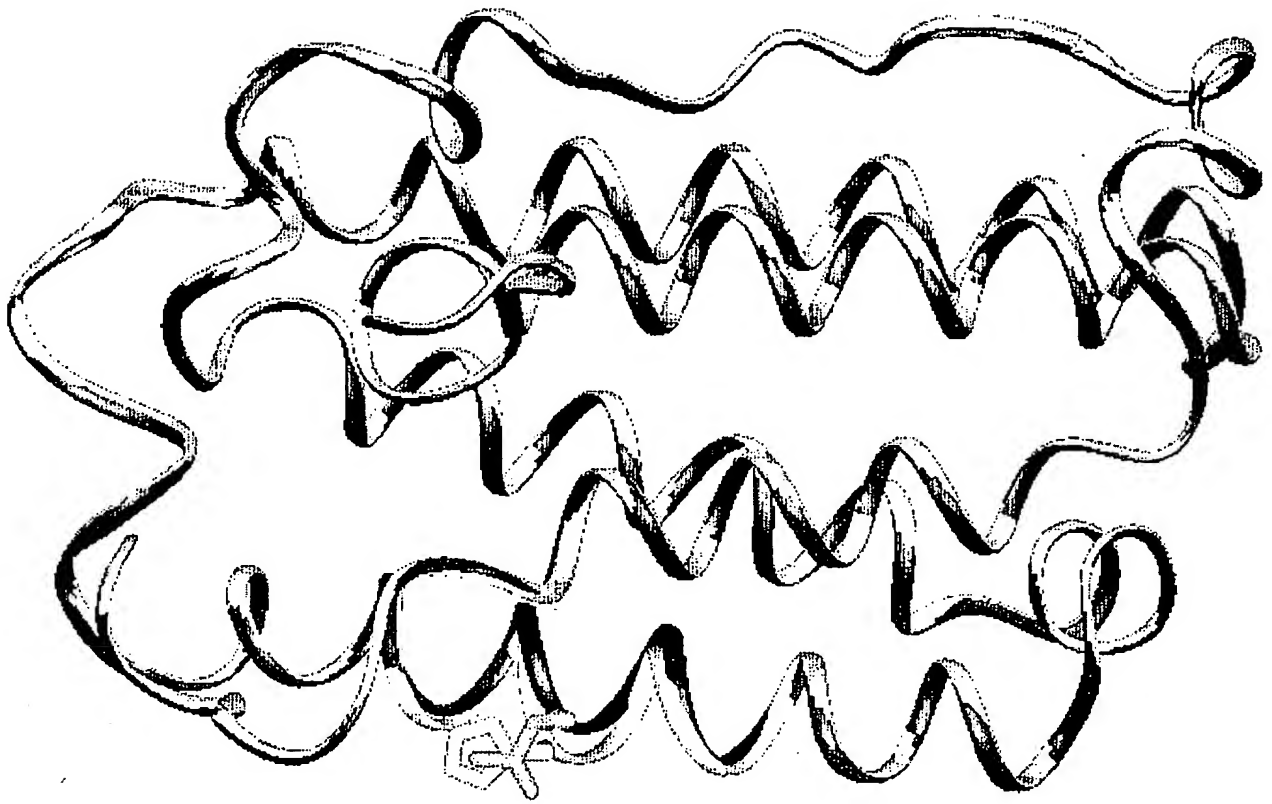


Figure 3

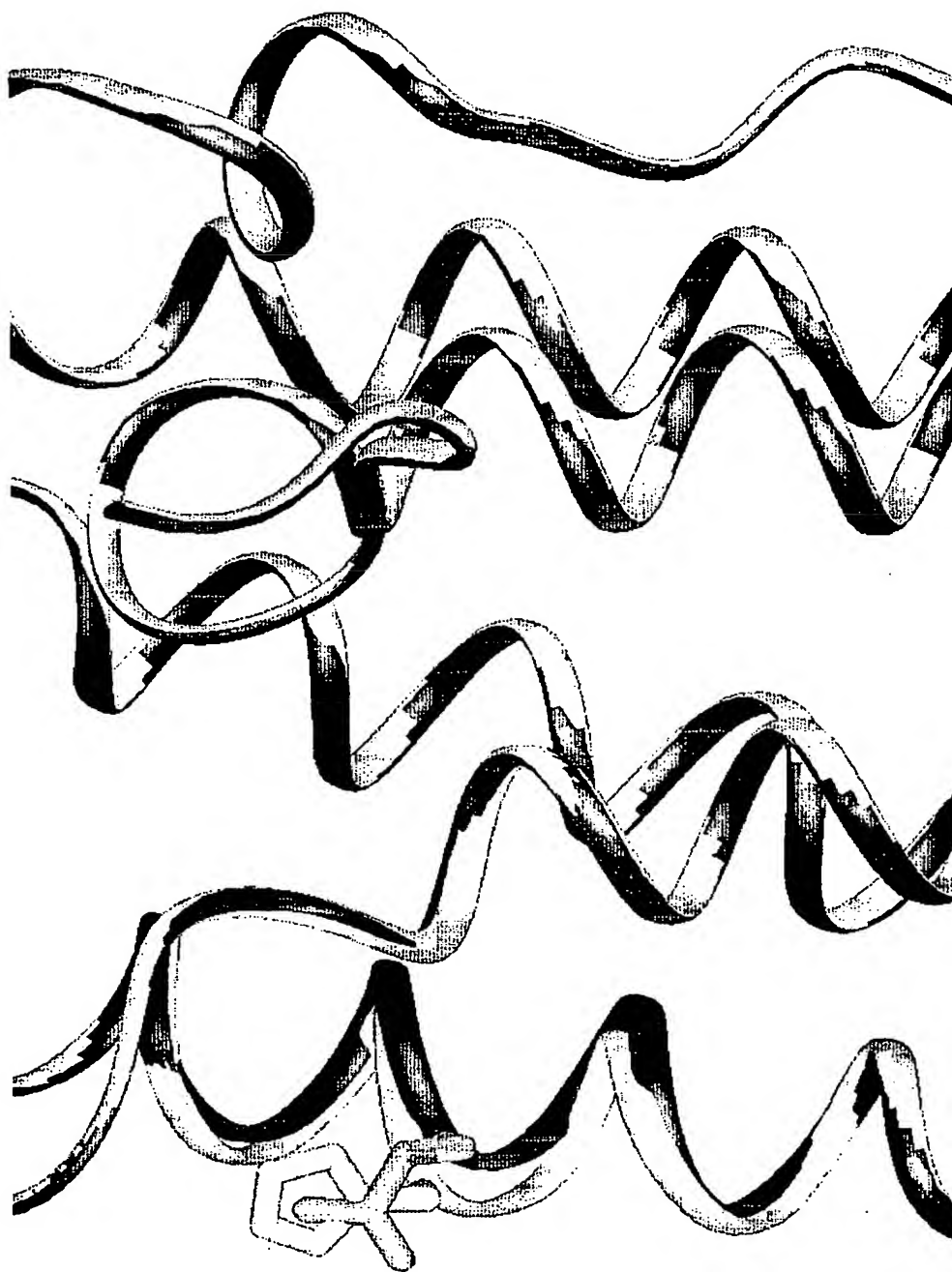


Figure 4

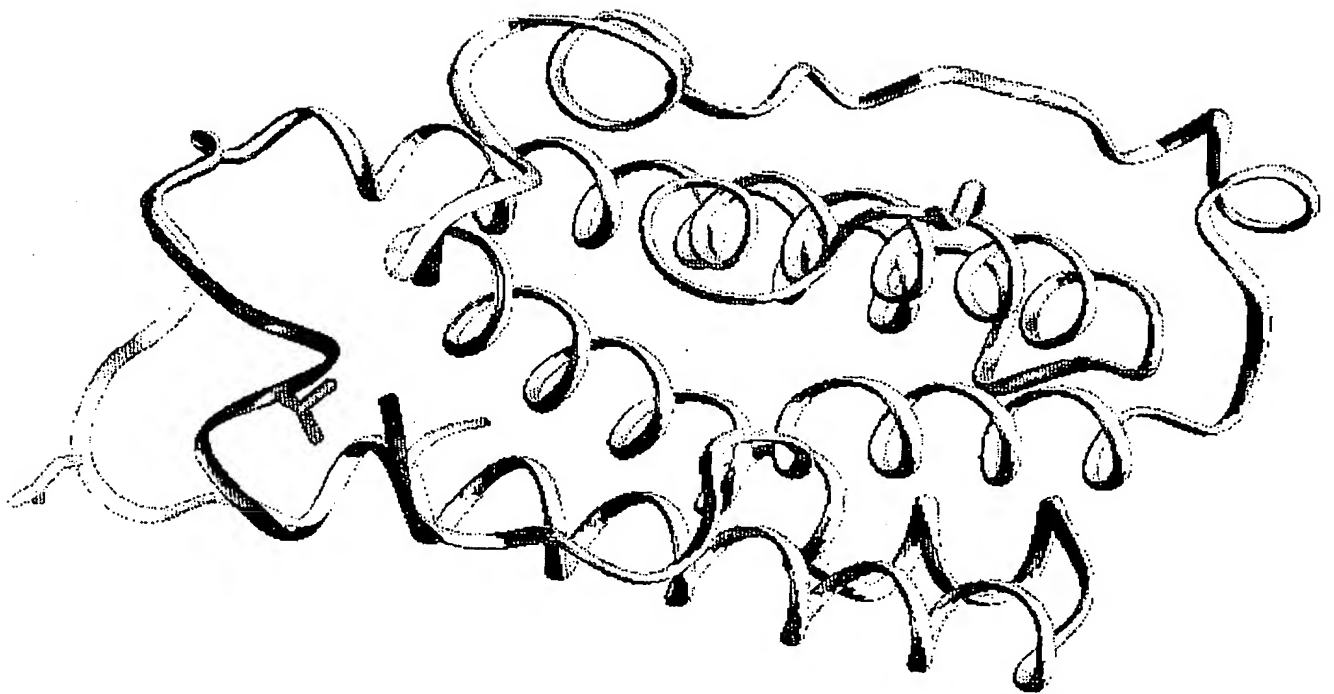


Figure 5

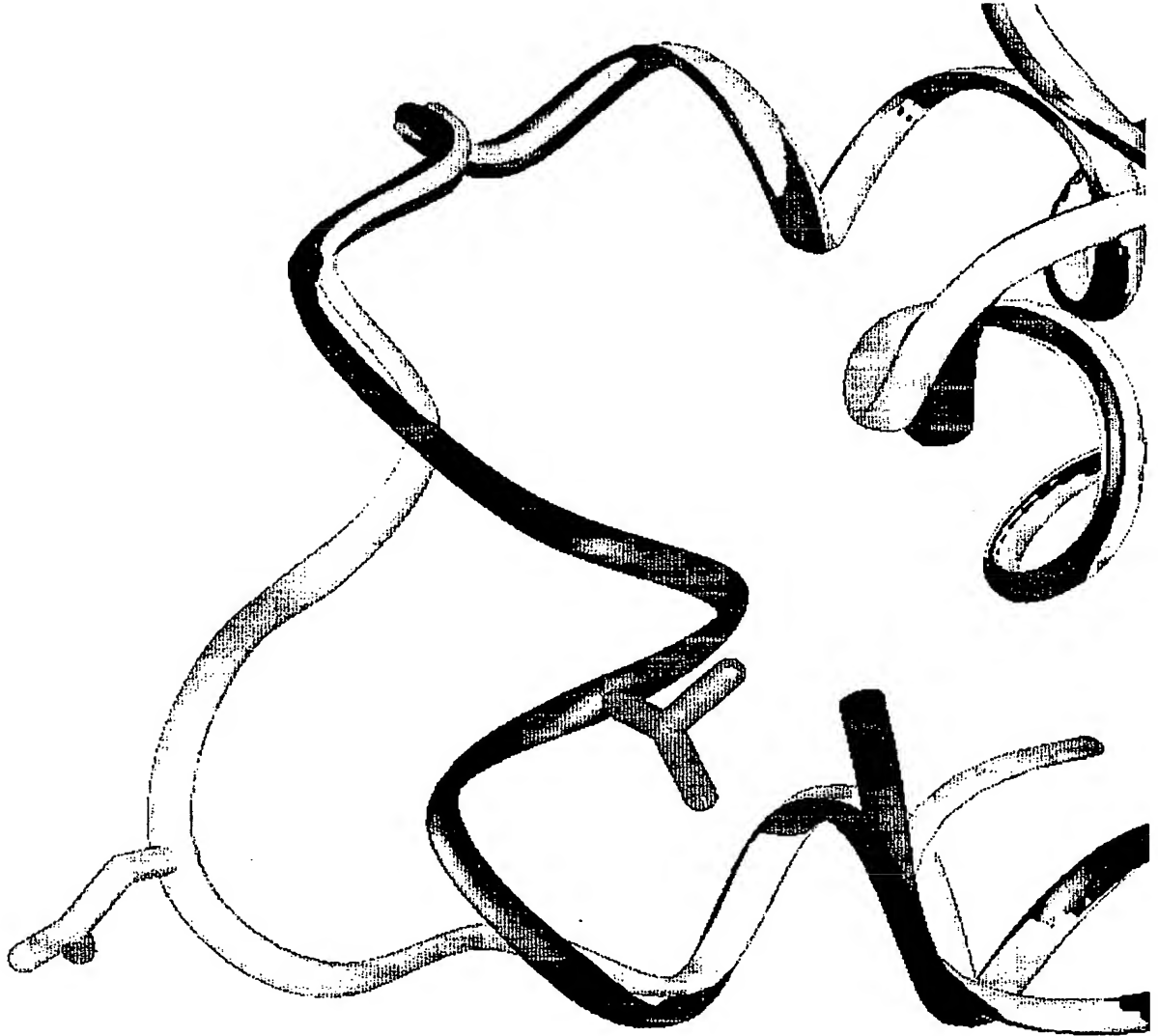
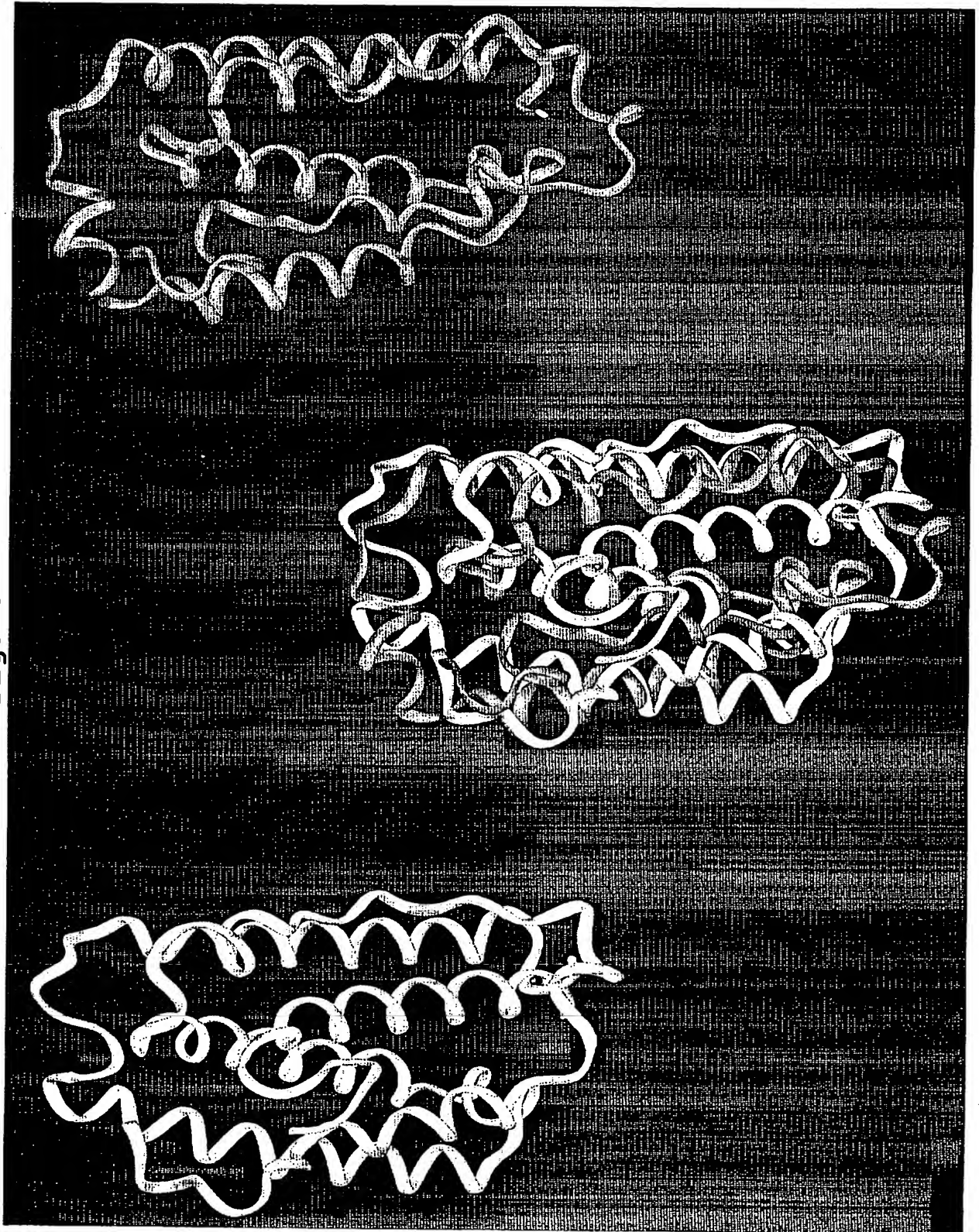


Figure 6

Figure 7



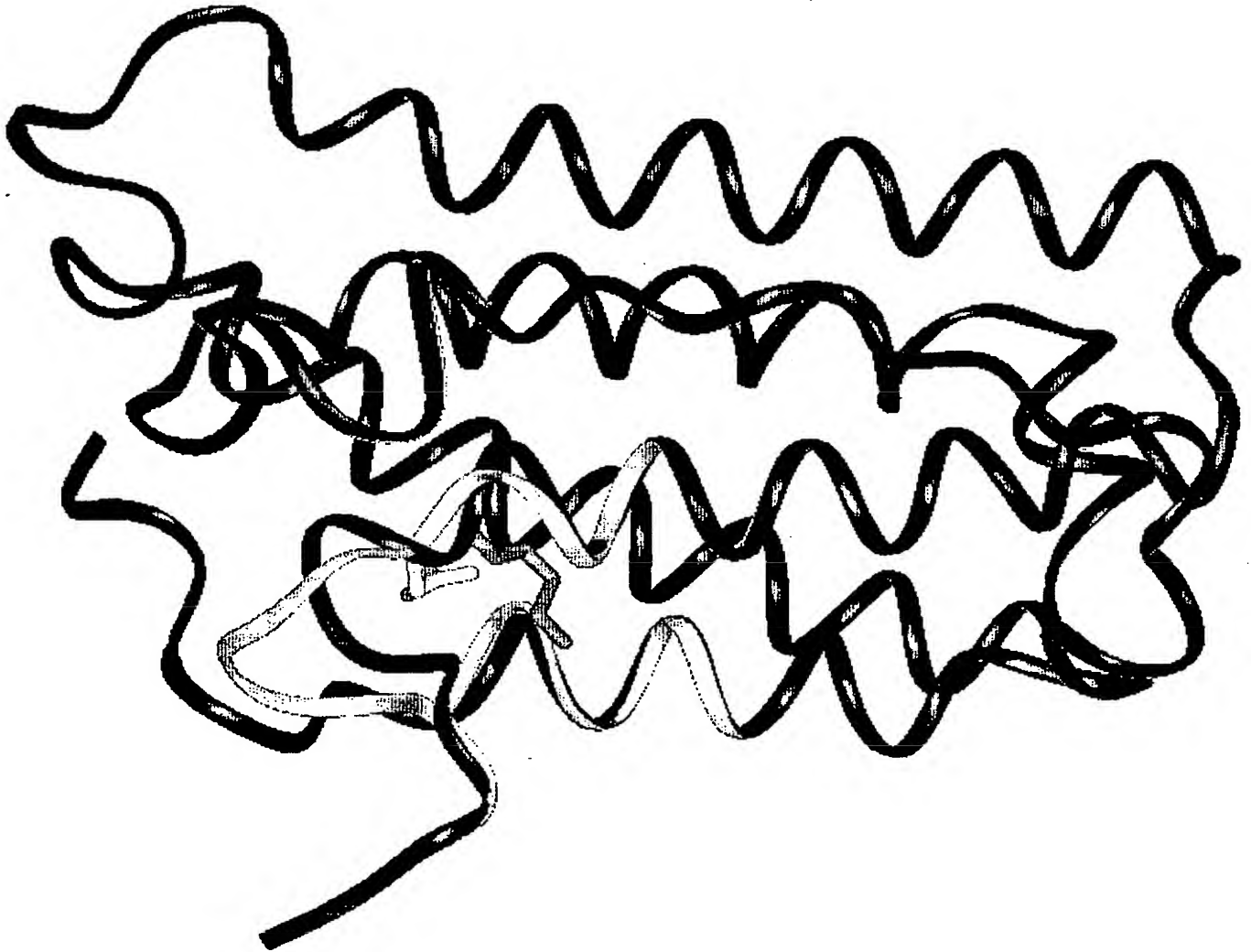


Figure 8

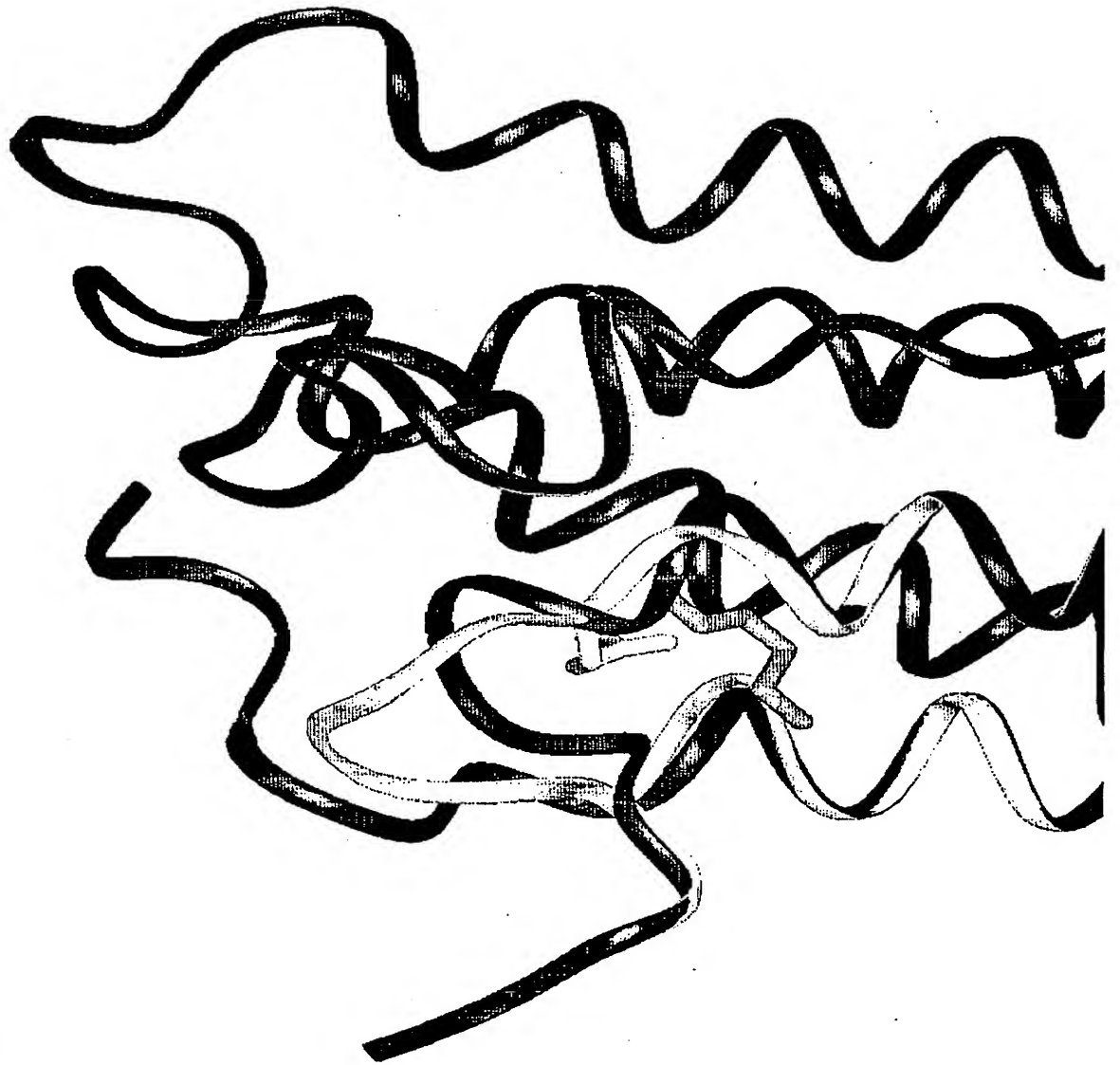


Figure 9

10/14

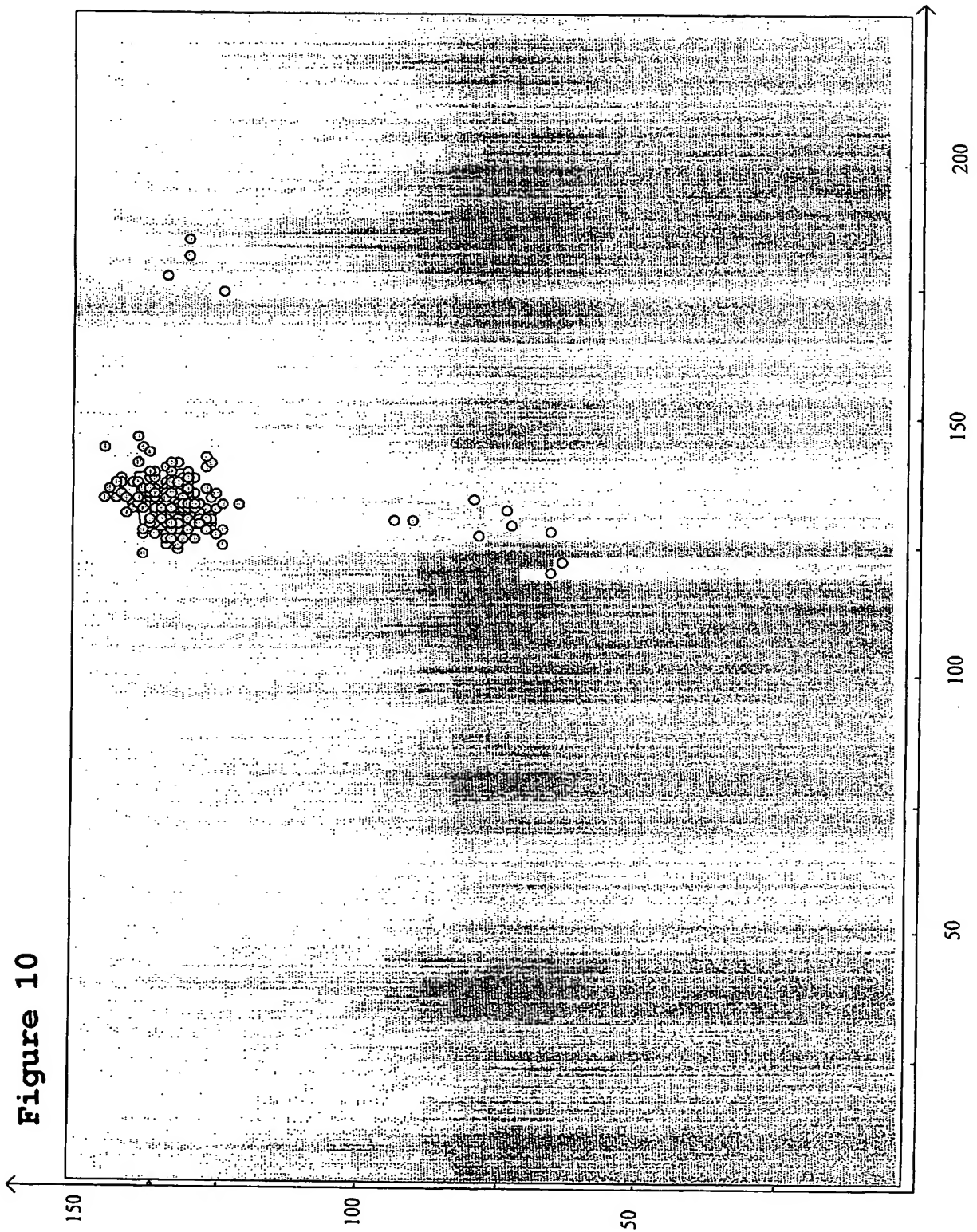


Figure 11

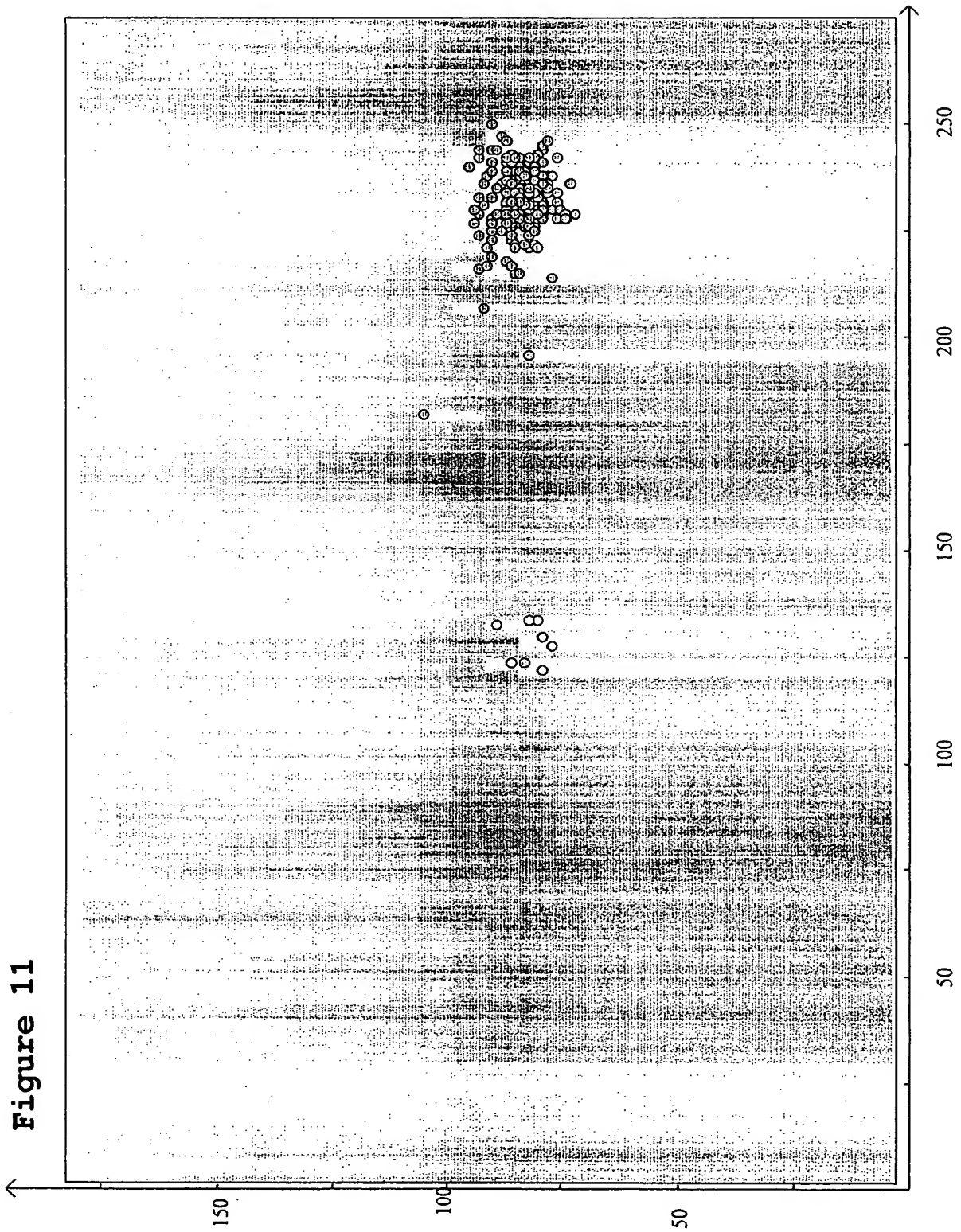


Figure 12

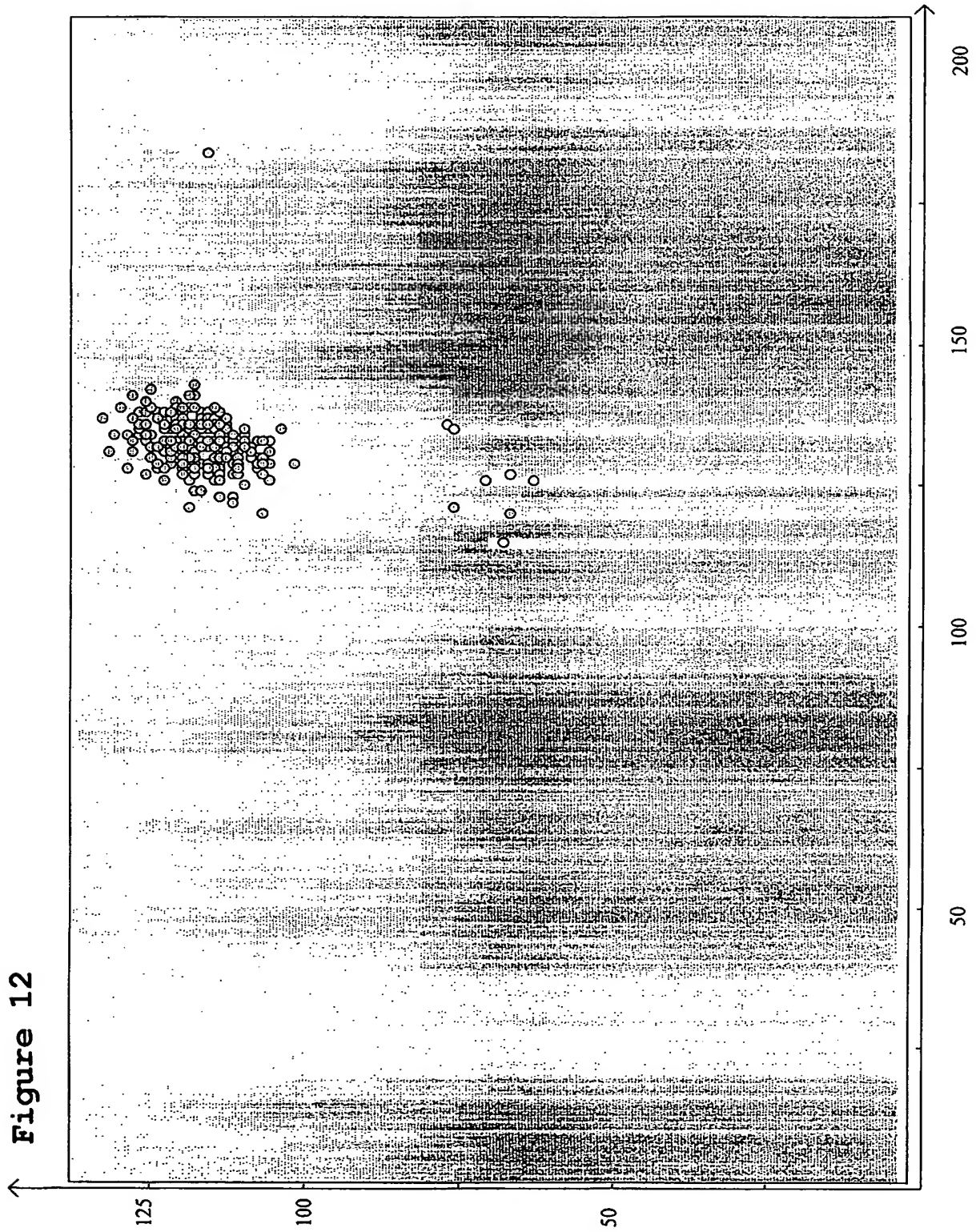
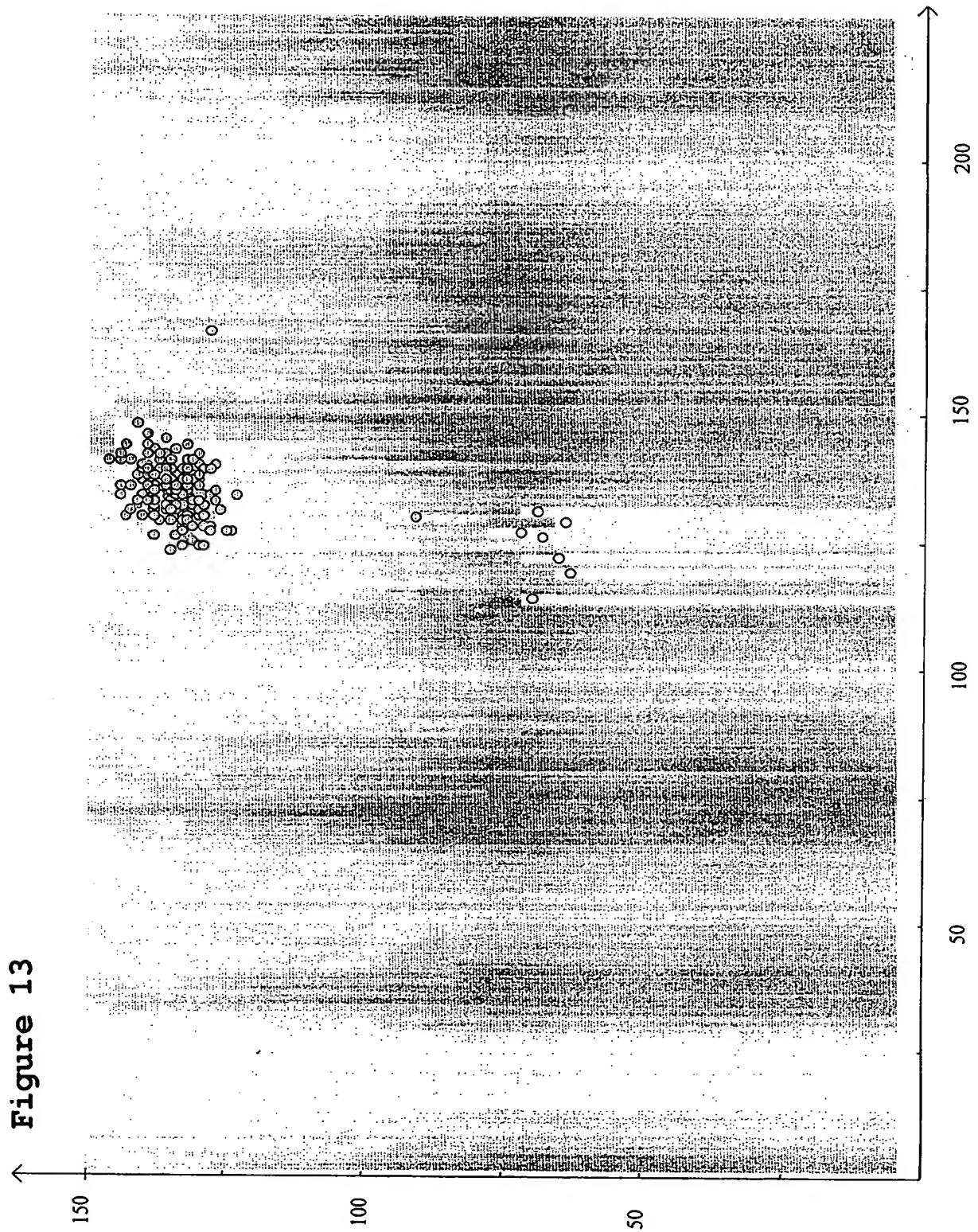
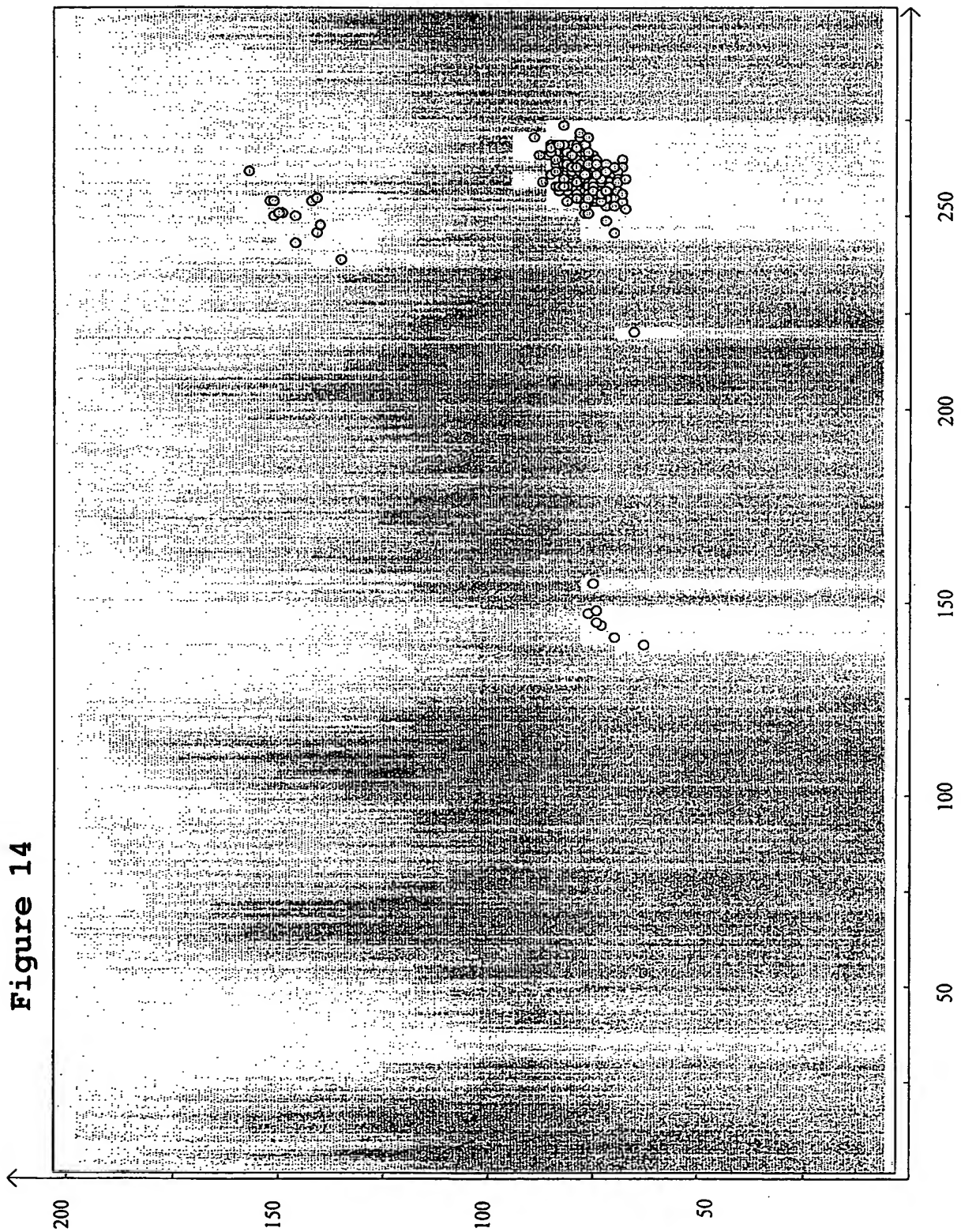


Figure 13





LISTE DE SEQUENCES

<110> GENODYSSEE

5 <120> Nouveaux polynucléotides comportant des polymorphismes
de type SNP fonctionnels dans la séquence nucléotidique
du gène IFNa-21 ainsi que de nouveaux polypeptides
codés par ces polynucléotides et leurs applications.

10 <130> IFNa21

<140>

<141>

15 <160> 4

<170> PatentIn Ver. 2.1

<210> 1

<211> 2001

20 <212> ADN

<213> Homo sapiens

<400> 1

attttgtcat ataaagcaaa attcaaagct tcatatatat actatgagaa aaatttttaa 60
25 aaattattga ttcataattt tagcagtttt gaatgattaa ctatgtaatt atattcatat 120
tattaatgtg tatttatata gatttttatt ttgcatatgt aatttcataa aacaaaattt 180
acatgaacaa attacattaa aagttattcc acaaataac ttatcaaatt aagttaaatg 240
tcaatagctt ttaaacttag attttagttt aacatttctg tcattcttta ctttgaataa 300
aaagagcaaa ctttatagtt tttatctgtg aagtagagat atacatatta tacataaata 360
30 gataagccaa atctgtgtta ctaaaatttc atgaagattt caattagaaa aaaataccat 420
aaaatgtttt gagtgcaggg gaaaaatagg caatgatgaa aaaaaatgaa aaacatctgt 480
aaacacatgt agagagtgc taaagaaagc aaaaacagag atagaaagta aaactagggc 540
atttagaaaa tggaaattag tatgttctact atttaagacc tacgcacaga gcaaagtctt 600
cagaaaacct agaggccaag gttcaagggt acccatctca agtagcctag caatattggc 660
35 aacatcccaa tggccctgtc cttttcttta ctgatggccg tgctggtgct cagctacaaa 720
tccatctgtt ctctgggctg tgatctgcct cagacccaca gcctgggtaa taggagggcc 780
ttgatactcc tggcacaaat gggaagaatc tctcctttct cctgcctgaa ggacagacat 840
gactttggat tccccagga ggagtttgat ggcaaccagt tccagaaggc tcaagccatc 900
tctgtcctcc atgagatgat ccagcagacc ttcaatctct tcagcacaaa ggactcatct 960

gctacttggg aacagagcct cctagaaaaa ttttccactg aacttaacca gcagctgaat 1020
 gacctggaag cctgcgtgat acaggagggtt ggggtggaag agactcccct gatgaatgtg 1080
 gactccatcc tggctgtgaa gaaatacttc caaagaatca ctctttatct gacagagaag 1140
 aaatacagcc cttgtgcctg ggaggttgtc agagcagaaa tcatgagatc cttctcttta 1200
 5 tcaaaaattt ttcaagaaag attaaggagg aaggaatgaa acctgtttca acatggaaat 1260
 gatctgtatt gactaataca ccagtcacac cttctatgac ttctgccatt tcaaagactc 1320
 atttctccta taaccaccgc atgagttgaa tcaaaaattt cagatctttt caggagtgtg 1380
 aggaaacatc atgtttacct gtgcaggcac tagtccttta cagatgacca tgctgataga 1440
 tctaattatc tatctattga aatattttatt tattttattag atttaaatta tttttgtcca 1500
 10 tgtaatatta tgtgtacttt tacattgtgt tatatcaaaa tatgtgattt atatttagtc 1560
 aatatattat ttcttttaat taaattttac tattaanaact tcttatatta ttgttttatt 1620
 ctttaataaa gaaataccaa gcccaaatgt gcaatctcat taaagaatgg atggtacaat 1680
 tcaattaccc atcatcattg tatccaaatt gtaagtaaaa attgactttc tctaagcgag 1740
 gttttatatt gcccttagga tatccagggtg aacataacaa ataccgtttt cgctttcttg 1800
 15 tatcttttat ttttgtaagg aaaataataa ctatactttc taatacctgt tacattaaat 1860
 gctatagtga gaagaaataa aaacaaatga aattcagtaa aactgaagca aggcataatca 1920
 aaattttttt taaaaaagta gtagatatcc tctatagcag acaagtagac atctaagtgc 1980
 aagtgtccat tggtaacctg a 2001

20 <210> 2
 <211> 189
 <212> PRT
 <213> Homo sapiens

25 <400> 2
 Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
 1 5 10 15

30 Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
 20 25 30

Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
 35 40 45

35 Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln Glu
 50 55 60

Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu
 65 70 75 80

[illegible]

```

    <210> 3
25  <211> 20
    <212> ADN
    <213> Homo sapiens

```

<400> 3
30 gggttcaagggt taccatctc

20

```

    <210> 4
    <211> 20
35  <212> ADN
    <213> Homo sapiens

```

```
<400> 4
tttgaaatgg cagaagtcac
```

20

**BREVET D'INVENTION****CERTIFICAT D'UTILITÉ**

Code de la propriété intellectuelle - Livre VI

N° 1123
DÉPARTEMENT DES BREVETS

26 bis, rue de Saint Pétersbourg

75800 Paris Cedex 08

Téléphone : 33 (1) 53 04 53 04 Télécopie : 33 (1) 42 94 86 54

DÉSIGNATION D'INVENTEUR(S) Page N° 1 / 1

(Si le demandeur n'est pas l'inventeur ou l'unique inventeur)

Cet imprimé est à remplir lisiblement à l'encre noire

DB 113 W /

| | | |
|---|----------------------|---|
| Vos références pour ce dossier (facultatif) | | BIF022965/FR |
| N° D'ENREGISTREMENT NATIONAL | | 0104404 |
| TITRE DE L'INVENTION (200 caractères ou espaces maximum) | | |
| Nouveaux polynucléotides comportant des polymorphismes de type SNP fonctionnels dans la séquence nucléotidique du gène IFNa-21 ainsi que de nouveaux polypeptides codés par ces polynucléotides et leurs utilisations thérapeutiques. | | |
| LE(S) DEMANDEUR(S) : | | |
| GenOdyssee | | |
| DESIGNE(NT) EN TANT QU'INVENTEUR(S) : (Indiquez en haut à droite «Page N° 1/1» S'il y a plus de trois inventeurs utilisez un formulaire identique et numérotez chaque page en indiquant le nombre total de pages). | | |
| Nom | | ESCARY |
| Prénoms | | Jean-Louis |
| Adresse | Rue | 4 rue Moxouris |
| | Code postal et ville | 78150 LE CHESNAY |
| Société d'appartenance (facultatif) | | |
| Nom | | |
| Prénoms | | |
| Adresse | Rue | |
| | Code postal et ville | |
| Société d'appartenance (facultatif) | | |
| Nom | | |
| Prénoms | | |
| Adresse | Rue | |
| | Code postal et ville | |
| Société d'appartenance (facultatif) | | |
| Nom | | |
| Prénoms | | |
| Adresse | Rue | |
| | Code postal et ville | |
| Société d'appartenance (facultatif) | | |
| DATE ET SIGNATURE(S) DU (DES) DEMANDEUR(S) OU DU MANDATAIRE (Nom et qualité du signataire) | | Le 10 octobre 2001 Thierry CAEN N°98.0600 RINUY, SANTARELLI |
| | | |